THE UNIVERSITY OF MELBOURNE

Resource Allocation in OFDM Cellular Networks

Thayaparan Thanabalasingham

Submitted in total fulfilment of
the requirements for the degree of

Doctor of Philosophy

Department of Electrical and Electronic Engineering
Faculty of Engineering
THE UNIVERSITY OF MELBOURNE
AUSTRALIA

December 2006

Produced in LaTeX 2ε

THE UNIVERSITY OF MELBOURNE
AUSTRALIA

Abstract

Resource Allocation in OFDM Cellular Networks

Thayaparan Thanabalasingham

The efficient use of radio resources is crucial in order for future wireless systems to be able to meet the demand for high speed data communication services. *Orthogonal Frequency Division Multiplexing* (OFDM) is an important technology for future wireless systems as it offers numerous advantages over other existing technologies, such as robust performance over multipath fading channels and the ability to achieve high spectral efficiency. Dynamic resource allocation can fully exploit the advantages of OFDM, especially in multiple user systems. In this thesis, we investigate a resource allocation problem in a multiple user, multiple cell OFDM cellular network focusing on downlink communications.

The majority of the resource allocation problems considered in the literature assume the availability of perfect channel state information at the base station for it to perform the resource allocation, which may not be realistic. Additionally, these problems are combinatorial in nature because of the discrete nature of the subcarrier allocation. On the other hand, we consider the case in which the base stations have only partial channel knowledge such as statistics of the channel variations. In this setting, it is not possible to find a resource allocation that will always meet the given user data rate requirements. Outage probability is an appropriate metric to measure the performance of the users.

Our resource allocation problem is to determine power and subcarrier allocations to ensure that the variations in the outage probabilities experienced by different users are minimal. We take a two-layer approach to solve this problem. The higher layer problem determines the power allocation for the base stations, and takes account of the coupling that exists between cells via the interference

that each base station provides to the others. Using the powers obtained by the higher layer allocation and exploiting the knowledge about the statistical variation in fading across subcarriers, the lower layer problem allocates the discrete subcarriers to the users in each cell to minimize the maximum outage probability in each cell. We make judicious use of interference averaging, continuous relaxation, and layering, and avoid a combinatorial explosion as the number of users and subcarriers grows large.

Declaration

This is to certify that

(i) the thesis comprises only my original work towards the PhD,

(ii) due acknowledgment has been made in the text to all other material used,

(iii) the thesis is less than 100,000 words in length, exclusive of tables, maps, bibliographies and appendices.

Signature_____

Date_____

Acknowledgments

I would like to express my gratitude to my supervisors for their cooperation and constant support. Firstly, I would like to thank A/Prof. Stephen Hanly and Dr. Lachlan Andrew for their continuous support and cooperation. They have taught me so much about doing research and their generosity and guidance are greatly appreciated. I would like to also thank Dr. Brian Krongold for being a co-supervisor during the final year of my candidature.

I would like to thank A/Prof. Subra Dey and A/Prof. Jamie Evans for serving in my Ph.D. committee. I am thankful to John Papandriopoulos for our collaboration, in particular on our ICC paper. I would also like to thank Liang Chen for several useful discussions.

I am grateful for the excellent research environment provided by the Australian Research Council (ARC) Special Research Centre for Ultra-Broadband Information Networks (CUBIN). CUBIN has been an ideal place for stimulating discussions with experts of many different fields. I also appreciate CUBIN for sponsoring my travel to various conferences and workshops. I would also like to thank my fellow CUBINites for many an enjoyable time.

Last but not least, I would like to thank my family. I am deeply grateful to my mother and my late father, who helped me to become what I am today. I would like to thank my mother and sisters for their encouragement to pursue this PhD.

To my mother and my late father

Contents

List of Figures

List of Tables

List of Acronyms

1G	First Generation
2G	Second Generation
3G	Third Generation
ADC	Analog to Digital Conversion
AMPS	Advanced Mobile Phone System
CDMA	Code Division Multiple Access
CSI	Channel State Information
DAC	Digital to Analog Conversion
DFT	Discrete Fourier Transform
DMT	Discrete Multiple Tone
DSL	Digital Subscriber Line
FDD	Frequency Division Duplex
FFT	Fast Fourier Transform
FRF	Frequency Reuse Factor
Gbps	Gigabits per second (10^9 bits per second)
GSM	Global System for Mobile Communications
ICI	Inter Carrier Interference
IDFT	Inverse Discrete Fourier Transform
IFFT	Inverse Fast Fourier Transform
IS-95	Interim Standard 95
ISI	Inter Symbol Interference
LLSA	Lower Layer Subchannel Allocation
LR	Lagrangian Relaxation
Mbps	Megabits per second (10^6 bits per second)
OFDM	Orthogonal Frequency Division Multiplexing
OFDMA	Orthogonal Frequency Division Multiple Access

PSD	Power Spectral Density
QAM	Quadrature Amplitude Modulation
QoS	Quality of Service
RNC	Radio Network Controller
SIR	Signal to Interference-and-noise Ratio
SNR	Signal to Noise Ratio
TACS	Total Access Cellular System
TDD	Time Division Duplex
UMTS	Universal Mobile Telecommunications System
WiMAX	Worldwide Interoperability for Microwave Access
WLAN	Wirless Local Area Network

Chapter 1

Introduction

1.1 Looking Forward

The history of wireless communications dates back more than a century [12]. The introduction of cellular telephone networks in the 1970s enabled the provision of wireless communications for a large number of mobile users within a geographical area while using a limited frequency spectrum [80]. The basic concept of cellular telephony is to divide the service area of the network into a number of small areas known as cells and reusing portions of available frequency spectrum in different parts of the network. Each cell is served by a base station transceiver.

The main focus of the first and second generation cellular systems was on voice traffic. The first generation (1G) cellular systems such as AMPS (Advanced Mobile Phone System) and TACS (Total Access Cellular System), deployed in the late 1970s and early 1980s, were analog in nature and supported voice services. The digital, second generation (2G) systems were introduced in the early 1990s and succeeded the first generation systems. These include GSM (Global System for Mobile Communications) and IS-95 (Interim Standard 95), and provide voice and low speed data services.

The widespread success of second generation cellular systems led to the development of newer wireless systems to support high speed data services. The third generation (3G) cellular systems, such as UMTS (Universal Mobile Telecommunications System) and CDMA2000 (Code Division Multiple Access 2000), started to be deployed at the beginning of the 21st century. They provide high speed data communications and offer content-based services such as wireless Internet and mobile TV [37, 80, 16, 45]. The turn of the century also saw the emergence

of other wireless technologies such as wireless local area networks (WLANs) [1] that provide fixed wireless access and complement the third generation cellular systems.

The future generation wireless systems are envisioned to provide ubiquitous communications to mobile users through cooperation between various radio access systems [6]. The future generation systems will achieve data rates of up to 1 Gbps, which are well above the maximum of 2 Mbps achieved by the third generation systems, and support data rate intensive multimedia applications such as high definition video [45, 65]. In order to be able to support the higher data rates demanded by future generation systems, the efficient use of radio resources is paramount.

Two of the most promising candidate technologies that are envisaged for future generation wireless communication systems are Code Division Multiple Access (CDMA) and Orthogonal Frequency Division Multiplexing (OFDM) [32]. In CDMA, all users share the available bandwidth and unique spreading codes are used to differentiate between users. The transmitted information is distributed across the available frequency spectrum by using these codes. Each user can be allocated one or more codes for data transmission. The resources that are allocated in a CDMA system are the spreading codes and transmit powers. Existing systems that use CDMA as the access technology include IS-95, CDMA2000 and UMTS.

OFDM is a multi-carrier transmission technology in which the available bandwidth is divided into a number of non-interfering (orthogonal) narrow band subcarriers. Each user can be allocated one or more subcarriers for data transmission. The resources allocated in an OFDM system are the subcarriers and the transmit powers. OFDM has been widely deployed in commercial systems, including in digital audio/video broadcasting [4, 3] and wireless LANs [1]. It has also been proposed as the access technology for future systems such as WiMAX [5].

In this thesis, we study the resource allocation for an OFDM cellular system in the presence of multipath fading.

1.2 Focus of the Thesis

This thesis studies dynamic resource allocation in a multiple user, multiple cell OFDM cellular system, and we focus on the downlink in which a base station is transmitting to mobiles. We consider the case that each user has a given data rate requirement.

Resource allocation in an OFDM context refers to allocating subcarriers to users and selecting the power levels and the modulation schemes on the allocated subcarriers, with the objective of meeting individual user quality of service (QoS) requirements, e.g., data rate requirements.

The majority of the resource allocation techniques in the literature work with the assumption that the transmitters (base stations, in the case of downlink communications) have knowledge of the instantaneous channel conditions, i.e., perfect channel state information (CSI), of all users in the system. The assumption of perfect CSI is not realistic, especially when the channel conditions change quickly with time, as in a mobile radio network. Thus, resource allocation can only be performed based on statistical knowledge of the channel conditions. In this setting, it is not possible to perform resource allocation in order to *guarantee* that the users receive the target data rates at all times; if a user is not able to receive reliably due to sufficient of its subcarriers fading simultaneously, then the user will be in outage. The outage probability is then the appropriate metric to measure the performance of the users.

We investigate the problem of determining the power and subcarrier allocations for users to ensure that users' outage probabilities are low.

Due to the difficulty associated with characterizing the outage probability as a function of all design parameters (number of subcarriers allocated to each user, transmit powers on each subcarrier, interference on each subcarrier, *etc.*), we use a two timescale, layered approach to solve this outage based resource allocation problem. We also employ fast frequency hopping based *interference averaging* to simplify the design. With interference averaging, the average interference to a user comes from all users in the neighbouring cells instead of coming from a

particular set of users. This removes the need to associate each subcarrier with a particular interferer, allowing us to model individual powers to each link instead of modeling individual powers to each subcarrier.

The higher layer problem addresses the issue of allocation of powers to users on a slow timescale and takes account of the coupling that exists between the cells created by the inter-cell interference. The fast fluctuations due to frequency selective fading and frequency hopping are averaged out in this closed-loop control formulation.

The lower layer resource allocation handles the allocation of the subcarriers on a faster timescale. By using the power allocation determined by the higher layer allocation, optimization of subcarriers for the users is performed in each cell without the need to keep track of the interference from other cells. This subcarrier allocation is performed based on the statistics of the channel conditions of the links, with the objective of *equalizing* the outages among the users.

1.3 Organization of the Thesis

The chapters in this thesis are organized as follows.

Chapter 2 is an overview chapter, which introduces the basics of OFDM. The chapter begins with a brief description of the signal propagation in wireless channels, introducing the concepts such as flat fading and frequency selective fading. This is followed by an overview of the OFDM technology. The basic operation of OFDM is described and the key features of OFDM that make it a robust modulation scheme are discussed. In addition, detrimental effects such as synchronization errors are also outlined.

Chapter 3 contextualizes our research problem by surveying the existing resource allocation techniques for OFDM systems in the literature. The chapter begins with surveying various adaptive resource allocation techniques for OFDM systems. An open area of research, namely, the problem of resource allocation in a multi-cell OFDM cellular network where the base stations only have statistical channel knowledge, is identified. Our resource allocation problem is defined

in Section 3.6 and a two-layer approach to solve it is outlined. Fast frequency hopping based interference averaging, which we employ in order to simplify the design, is also introduced.

Chapters 4 and 5 are dedicated to solving the higher layer resource allocation problem which is equivalent to a resource allocation problem in a flat fading environment. The system model is developed and the higher layer resource allocation problem is formulated as a power minimization problem. A distributed algorithm is devised to solve this problem. A corresponding discrete version of the problem is also solved.

Chapter 5 is about reducing the complexity of the higher layer resource allocation. This is achieved by imposing additional constraints on the resource allocation problem formulated in Chapter 4. Two such reduced complexity schemes are considered. The first scheme is the static subcarrier allocation scheme in which the allocation of subcarriers are done a priori and appropriate power allocation for the users is then selected. The second scheme is the uniform power spectral density (PSD) allocation scheme which constrains the transmitter (base station) to use equal power on each subcarrier. Distributed algorithms are developed to solve these problems. These reduced complexity schemes are numerically compared with the exact solution (Chapter 4) to quantify the loss in performance due to the reduction in the degrees of freedom. The corresponding discrete versions of these reduced complexity problems are also solved.

Chapter 6 deals with the resource allocation in a multipath frequency selective environment. Focusing on the uniform PSD allocation scheme as the method of choice at the higher layer, the lower layer allocation problem is formulated to determine the precise allocation of subcarriers to users by taking account of the statistical knowledge about the subcarriers. An algorithm, which operates locally in each cell, is developed to allocate subcarriers to users within the cell in order to "balance" the outage probabilities. The balancing of outage probabilities refers to minimizing the maximum outage probability of the users. Monte Carlo simulations are conducted to evaluate the performance of the algorithm.

1.4 Contributions of the Thesis

The ultimate objective of this thesis is to solve a downlink resource allocation problem in an OFDM cellular network under the influence of frequency selective fading, with imperfect channel state information. A two layer approach is used to solve this problem (as outlined in Section 3.6). The main contributions of this thesis are as follows.

1.4.1 Chapter 4

The higher layer resource allocation problem which is equivalent to a problem in a flat fading environment is considered in this chapter. Allowing a continuous allocation of subcarriers (i.e., bandwidth) to users, the resource allocation problem is formulated as a power minimization problem which explicitly models the coupling between cells that exists due to the inter-cell interference. It is shown that although this problem is not convex, it can be transformed into one by appropriate transformations. Furthermore, it is demonstrated that solving this transformed problem does not lead to a distributed solution.

A distributed algorithm based on Yates framework [106] is proposed to solve the problem. Using a similar approach, the discrete version of the higher layer problem (where the allocation of subcarriers to users is discrete) is also solved using a distributed algorithm.

Part of the work in this chapter was presented in [94].

1.4.2 Chapter 5

Continuing the analysis of the higher layer problem which was formulated in Chapter 4, two variants to this problem, namely, a static bandwidth allocation and a uniform power spectral density (PSD) allocation, are considered in this chapter by imposing additional constraints. A distributed algorithm based on Yates framework [106] is devised to solve the static bandwidth allocation problem.

The uniform PSD allocation scheme is extensively studied. By using the Yates framework [106], the properties of the optimal solution of the uniform PSD allocation scheme are characterized and a distributed algorithm is developed. As the algorithm based on Yates framework consists of iterations within iterations to reach the optimal solution, a new algorithm which only requires a single iteration is constructed. As this new algorithm does not fall under Yates framework, its convergence is proved using a novel technique based on monotonicity.

Using a simulation study, it is shown that the loss in performance in either of the reduced complexity schemes is small compared to the scheme that solves the higher layer resource allocation problem exactly, even though the number of parameters available for the reduced complexity problems are about a half of that of the original problem. Of the two reduced complexity schemes, the uniform PSD allocation scheme slightly outperforms the static bandwidth allocation scheme.

The discrete version of the uniform PSD allocation scheme is also studied and a distributed algorithm is devised to solve it.

Part of the work in this chapter was presented in [93] and [94].

1.4.3 Chapter 6

This chapter provides a two layer approach to solve the outage probability based resource allocation problem in a multipath frequency selective fading environment. Using the uniform PSD allocation as the scheme of choice at the higher layer and fixing the powers to the values obtained by the higher layer, the lower layer resource allocation problem is formulated as a single cell problem with the objective of minimizing the maximum outage probability which takes account of statistical knowledge about the subcarriers to allocate discrete subcarriers to users. As solving the resulting combinatorial problem is numerically complex, this problem is approximated by another, simpler combinatorial problem using the central limit theorem.

A lower layer subcarrier allocation (LLSA) algorithm is developed to solve this approximate problem exactly in two stages: the first is a continuous relax-

ation, and the second refines the relaxed solution to the optimal discrete solution in at most as many steps as the number of users in the cell. It is shown by simulations that the LLSA algorithm works well in balancing the outages among the users.

Part of the work in this chapter forms the basis of [92].

Chapter 2

Background

2.1 Introduction

Orthogonal Frequency Division Multiplexing (OFDM) can be thought of as a modulation technique as well as a multiple access scheme. As a modulation scheme, it is well suited to handle adverse environmental conditions, while as a multiple access scheme, it offers high spectral efficiency and diversity.

OFDM is a special form of multi-carrier transmission scheme. It offers better spectral efficiency over the traditional multi-carrier systems. The basic idea of OFDM is to split a high rate data stream into a large number of of lower rate data substreams and modulate them onto a number of specially designed carrier frequencies called *subcarriers*. The data is transmitted simultaneously over these parallel subcarriers.

The history of OFDM dates back to the mid 1960s [22, 84, 23]. Early implementations of OFDM were however limited to military applications due to the difficulties involved with its implementation. Only the recent advances in integrated circuit technology enabled its cost effective implementation, paving the way for the development of commercial applications, in both wired and wireless environments. When it is used in the wired environment, it is typically referred to as *discrete multiple tone* (DMT) [88].

The focus of this thesis is on the wireless environment. The wireless channel is much more unpredictable than the wireline channel because of the factors such as shadowing, the multipath effect and the Doppler effect. These factors impact how the signal level changes as it propagates through wireless medium. OFDM offers greater immunity to the impairments and the uncertainties of the wireless

channel.

This chapter presents essential background information on OFDM that will provide the foundation for the work carried out in later chapters. Section 2.2 briefly reviews the signal propagation in the wireless environment. Various key principles of OFDM are described in Section 2.3.

2.2 Wireless Propagation

The radio signals are electromagnetic waves and the propagation of electromagnetic waves through a transmission medium is generally governed by three natural phenomena: reflection, diffraction and scattering [80]. They cause variations in the strength of the waves as the waves propagate. Depending on the time scale these variations occur, they can be grouped into two categories: large scale fading (slowly varying in time and space) and small scale fading.

2.2.1 Large Scale Fading

The large scale fading refers to the degradation of the signal strength due to the path loss as a function of the distance between the transmitter and the receiver and shadowing effects caused by the surrounding environmental clutter. The path loss is the gradual loss of received signal power with the distance from the transmitter. Shadowing describes the random effects which occur at different locations which have the same transmitter receiver separation, but different surroundings on the propagation path.

Path Loss

When the transmitter and the receiver have a clear unobstructed line-of-sight path between them, the free space propagation model (Friis model [80]) can be used to predict the received signal strength: the received signal strength P_r has a inverse-square relationship with the transmitter-receiver separation d, i.e., $P_r \propto d^{-2}$. However, the Friis model is only valid for distances d which are in the *far-*

field of the transmit antenna. The far-field of the transmit antenna is defined as a distance beyond a far-distance d_f which is a function of the largest linear dimension of the transmit antenna aperture and the carrier wavelength. For this reason, a known reference distance $d_0 \geq d_f$ is defined and the received power at any distance $d \geq d_0$ can be computed as:

$$P_r(d) = P_r(d_0)\left(\frac{d_0}{d}\right)^2, \qquad d \geq d_0,$$

where $P_r(d_0)$ is the received power at the reference distance d_0.

When there is no unobstructed line-of-sight path between the transmitter and the receiver, the prediction of the Friis model is far too optimistic. Both analytical and empirical models indicate that the average received signal power decreases according to:

$$P_r(d) = P_r(d_0)\left(\frac{d_0}{d}\right)^{\gamma}, \tag{2.1}$$

where γ is the *path loss exponent* which indicates the rate at which path loss increases with distance d. The value of γ depends on the specific propagation environment and typically ranges from 2 (i.e., free space attenuation) to 6 [80].

Shadowing

Equation (2.1) only gives an ensemble average value, which does not take into account the variability in the propagation path between the transmitter and the receiver. Empirical data measurements have shown that the received signal power at a particular distance d is distributed log-normally (normally in dB) about the distance-dependent average received signal power. When the log-normal shadowing is included, the large scale fading at a distance d, can be expressed as:

$$P_r(d) = P_r(d_0)\left(\frac{d_0}{d}\right)^{\gamma} 10^X \tag{2.2}$$

where X represents a normally distributed random variable with zero mean and variance σ^2. The variance of the log-normal shadowing is calculated based on

measurements that are taken over a wide range of locations that have the same transmitter receiver separation but with different environmental clutter.

Apart from this simple model of (2.2), there exist various empirical models that take into account the *terrain profile* of a particular area in the propagation predictions (such as Okumura model [66] and Hata model [40]).

2.2.2 Small Scale Fading

The small scale fading refers to the rapid fluctuations in the envelope of the received signal over very small distances or time periods. These are caused by multipath delay effects and mobility in the wireless environment.

Multipath Effect

The transmitted signal in a wireless environment may follow multiple paths to reach the receiver. As different replicas of the signal travel in different paths, they may arrive at the receiver at different times with widely varying amplitude and phase. This causes rapid fluctuations in the received signal strength due to the way these replicas of the signal add up. There are two related parameters that are used to describe the multipath delay effects. The *multipath delay spread* represents the difference in time between the earliest and the latest replica of the signal to arrive at the receiver. The coherence bandwidth which is inversely proportional to the delay spread, provides a measure of the range of frequencies over which the channel can be considered "flat".

When the bandwidth of the transmitted signal is smaller than the coherence bandwidth of the wireless channel, then the signal will experience *flat fading*. Flat fading channels are also known as *narrow band* channels, since the signal bandwidth is narrow compared to the coherence bandwidth of the channel. In the time domain, this corresponds to the symbol period of the signal being larger than the delay spread of the signal.

On the other hand, *frequency selective fading* occurs when the bandwidth of the transmitted signal is larger than the coherence bandwidth of the wireless channel.

In this case, the signal will have frequency components that are spaced apart by more than the coherence bandwidth and these components will experience different attenuations. Frequency selective fading channels are also known as *wide band* channels. In the time domain, this corresponds to the symbol period of the signal being smaller than the delay spread of the signal. As a result, the signal components of different symbols overlap introducing what is known as *inter-symbol interference* (ISI).

Doppler Effect

The Doppler effect captures the frequency dispersion caused by the relative motion between the transmitter and the receiver or by the movements of the objects surrounding them. This frequency dispersion is an apparent shift in frequency of the received signal, and is referred to as Doppler Shift. The Doppler Shift is directly proportional to the velocity and the direction of the relative motion of the receiver with respect to the transmitter [80]. There are two related parameters that are used to describe the Doppler effect: Doppler spread and the coherence time. The Doppler spread characterizes the spectral broadening of the signal, and is defined as the range of frequencies over which the received spectrum of the carrier signal is non-zero. The coherence time is inversely proportional to the maximum Doppler shift, and is a measure of the time invariance of the channel. The coherence time represents the period over which two received signals have a strong amplitude correlation.

When the Doppler spread is high relative to the signal bandwidth, the signal experiences *fast fading*. Fast fading causes the channel to fluctuate faster than the variations in the signal. In the time domain, this corresponds to the symbol period of the signal being larger than the coherence time of the channel, which causes significant distortion of the symbol. The spectrum broadening due to the Doppler shift cannot be ignored at the receiver.

Conversely, if the signal bandwidth is much greater than the Doppler spread, then the effects of the Doppler shift is negligible at the receiver. This corresponds to a *slow fading* channel. The channel variation is much slower than the symbol

period of the signal, and the signal is not distorted to a significant degree. In the time domain, this corresponds to the coherence time of the channel being much larger the symbol period of the signal.

2.3 OFDM Technology

We will discuss key features of OFDM technology in this section.

2.3.1 Parallel Transmission

In wideband wireless systems, the bandwidth of the channel is much larger than its coherence bandwidth and thus the channel is frequency selective. With a single carrier transmission system, the symbols are transmitted sequentially, and the signals occupy the entire available bandwidth. Therefore, adaptive equalization techniques [79] are essential at the receiver for successful decoding of information. The equalizer estimates the channel impulse response to compensate for the distortion due to the multipath fading. As the symbol period becomes shorter with increasing data rates, the adaptive equalization becomes more difficult. The shorter the symbol duration, the more the number of adjacent symbols affected by a single fade will be.

On the other hand, in a multi-carrier system, the entire available bandwidth is divided into many narrow band subcarriers and each subcarrier experiences flat fading. Equalization is performed on a subcarrier basis and thus is much simpler than the single carrier transmission systems [15]. Fading across different subcarriers provides frequency diversity. A single fade only affects a limited number of the subcarriers. By coding across the symbols in different subcarriers, the information conveyed by those affected subcarriers can be recovered at the receiver [33].

The approach of parallel transmission also reduces the sensitivity of the system to the multipath delay spread. Since each subcarrier carries only a small fraction of the overall data rate, the symbol period of the signal is increased. The

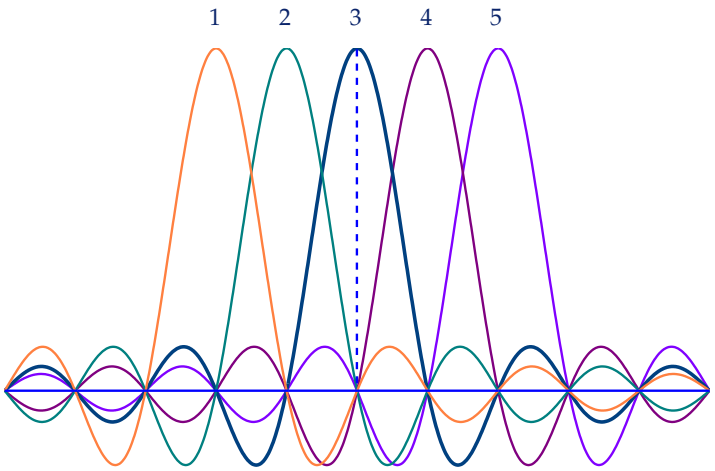


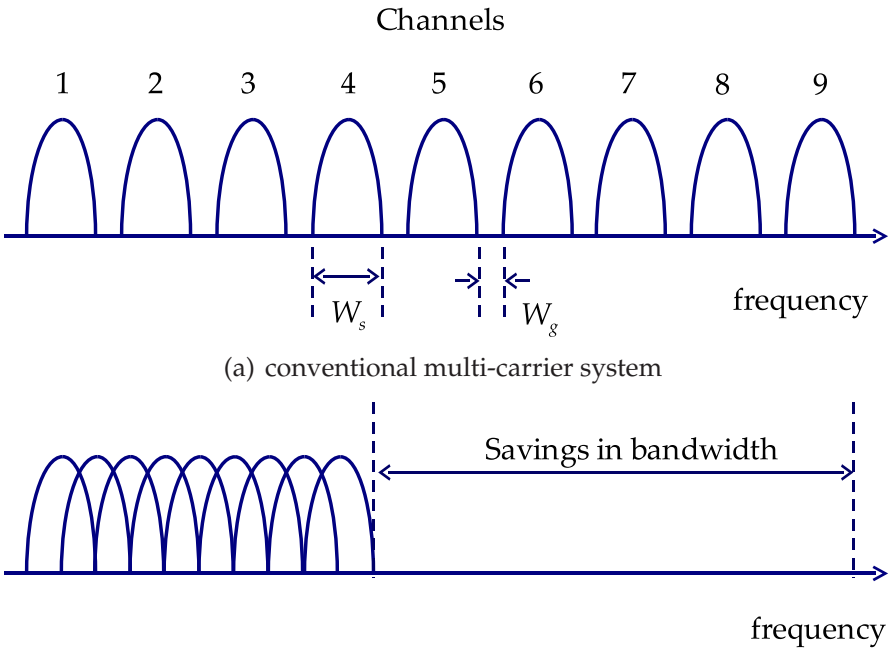
Figure 2.1: Example of an OFDM signal spectra: the spectrum of each subcarrier has a null at the center frequency of each of the others.

increased symbol period improves the robustness of OFDM to delay spread. This in effect reduces the effect of ISI. By introducing a guard interval called *cyclic prefix* [72], which we describe in Section 2.3.6, the ISI can further be reduced or even eliminated completely.

2.3.2 Orthogonality of Subcarriers

Another advantage of OFDM is the efficient use of the available frequency spectrum. In a conventional multi-carrier system, the frequency band is divided into non-overlapping subcarriers in order to eliminate the cross-talk between subcarriers known as *inter-carrier interference* (ICI). This non-overlapping design of the subcarriers leads to inefficient use of the available spectrum.

On the other hand, the OFDM offers high spectral efficiency by allowing the overlapping of the spectrum of the subcarriers. In order for this to work, the ICI between subcarriers must be mitigated. This is achieved by making the subcarriers mutually orthogonal. The orthogonality between subcarriers is maintained by carefully selecting the spacing between the subcarriers. The orthogonality of the subcarriers means that each subcarrier has an integral number of cycles over



(a) conventional multi-carrier system

(b) OFDM system

Figure 2.2: Spectral efficiency of OFDM [98]: Both systems have the same subcarrier bandwidth W_s. The conventional system uses a guard band W_g between adjacent subcarriers. On the other hand, the spectra of the OFDM subcarriers overlap, leading to higher spectral efficiency.

a symbol period. Consequently, there is a difference of an integral number of cycles between any two subcarriers over a symbol period. This ensures that the spectrum of each subcarrier has a null at the center frequency of each of the other subcarriers in the system [103]. Figure 2.1 shows an example of an OFDM signal spectra. With perfect synchronization at the receiver, the information on each subcarrier could be decoded successfully without the interference from other subcarriers.

Figure 2.2 shows the savings of bandwidth with OFDM compared to a conventional system with the same subcarrier bandwidth. A saving of almost 50% of bandwidth can be achieved with OFDM due to the overlapping of subcarriers.

The orthogonality between subcarriers corresponds to a precise mathematical relationship between the subcarrier signals over a symbol period T_s. Suppose that the set $\{\Psi_k(t)\}$ represent the subcarriers signals, with $k = 0, \ldots, N_c - 1$. Then, orthogonality between subcarriers corresponds to the condition that over a symbol

period T_s, the signal contribution to a subcarrier k from all other subcarriers is zero:

$$\int_0^{T_s} \Psi_k^*(t)\ \Psi_l(t)dt = A \times \delta[k-l] = \begin{cases} A, & \text{if } k = l \\ 0, & \text{otherwise,} \end{cases} \tag{2.3}$$

where "$*$" is the complex conjugate operator, A is a constant, and $\delta[\cdot]$ is the Kronecker delta function.

Orthogonal sets of subcarrier signals can be obtained by using the sinusoidal wave family and also letting the spacing between subcarriers, Δf, equal to the reciprocal of the symbol period: $\Delta f = 1/T_s$. The orthogonal signals $\{\Psi_k(t)\}$ are of the form:

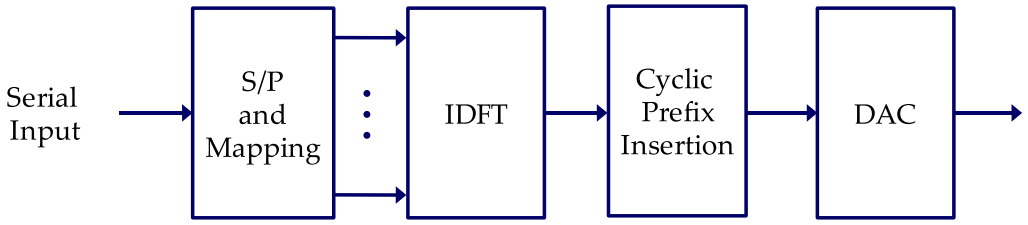
$$\Psi_k(t) = \begin{cases} e^{j2\pi f_k t}, & \text{if } t \in [0, T_s] \\ 0, & \text{otherwise,} \end{cases} \tag{2.4}$$

with f_k being the center frequency of subcarrier $k = 0, \ldots, N_c - 1$. It can easily be verified that the subcarrier signals of (2.4) satisfy the orthogonality condition of (2.3).

OFDM systems are more susceptible than single carrier systems to the effects of frequency offset errors due to the mismatch in the carrier frequencies in the transmitter and the receiver, and the Doppler effect due to mobility. Frequency offset errors cause the orthogonality between the subcarriers to be destroyed. Frequency synchronization addresses these imperfections as discussed in Section 2.3.8.

2.3.3 OFDM Transmission

A schematic diagram of an OFDM transmitter is depicted in Figure 2.3. The high rate data stream is divided into blocks of bit streams and each block contributes to an OFDM symbol. The block of bit stream is further split into N_c bins. The number of bits in each bin depends on the modulation scheme chosen for the corresponding subcarrier. The choice of the modulation scheme is in turn determined



S/P – serial to parallel conversion
DAC – digital to analog conversion

Figure 2.3: OFDM transmitter

by the quality of the subcarriers. The modulation scheme can vary between subcarriers.

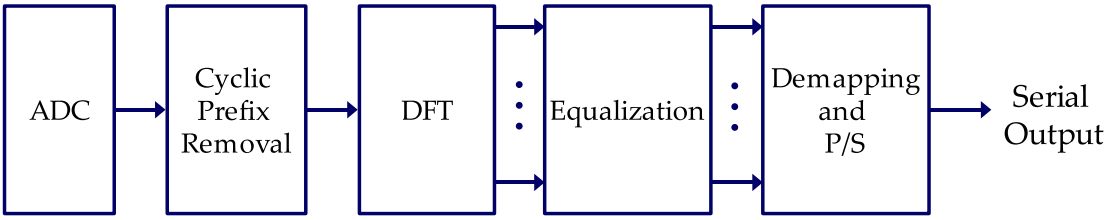
The data symbol for each subcarrier is selected based on the modulation scheme chosen. Let x_k be the data symbol selected for subcarrier k. The symbols $\{x_k\}$ are modulated onto the subcarriers using *inverse discrete Fourier transform* (IDFT). The IDFT operation essentially is a transformation of the OFDM symbol from the frequency domain to time domain. The output of this operation is a vector of time samples $\{S_m\}$ where:

$$S_m = \frac{1}{\sqrt{N_c}} \sum_{k=0}^{N_c-1} x_k e^{j2\pi k \frac{m}{N_c}}, \qquad m = 0, \ldots, N_c - 1.$$

After the IDFT operation, a cyclic prefix is added to the OFDM symbol prior to digital to analog conversion (DAC). The DAC output is a base band analog signal which is then up-converted and transmitted.

2.3.4 OFDM Reception

Figure 2.4 shows a typical OFDM receiver. At the receiver, the received signal is down-converted to base band, and analog to digital conversion (ADC) is performed. After the removal of the cyclic prefix, the time samples $\{R_m\}$ are fed into



P/S – parallel to serial conversion
ADC – analog to digital conversion

Figure 2.4: OFDM receiver

the DFT module and the output is a vector $\{y_k\}$ with:

$$y_k = \frac{1}{N_c} \sum_{m=0}^{N_c-1} R_m e^{-j2\pi k \frac{k}{N_c}}, \qquad k = 0, \ldots, N_c - 1.$$

Note that y_k's are the scaled versions of the transmitted data symbols x_k's. The channel response estimates are used to compensate for the amplitude attenuations and phase shifts in the y_k's. Since the frequency responses of the subcarriers are relatively flat, a simple one tap equalization is sufficient for detecting the data constellations. The estimate of the transmitted symbol is:

$$\hat{x}_k = \frac{y_k}{\hat{H}_k}, \qquad k = 0, \ldots, N_c - 1,$$

where $\hat{H}_k$ is the channel response estimate of the frequency band of subcarrier k.

2.3.5 IDFT/DFT

When OFDM was first proposed, the idea was to use arrays of sinusoidal generators and filters for performing multi-carrier modulation and demodulation. For a system with a large number of subcarriers, this became unreasonably complex. Because of this, earlier OFDM systems such as KATHRYN [111] used a small number of low rate channels. It was soon realized that the multi-carrier modulation and demodulation can be realized by a series of Fourier transform

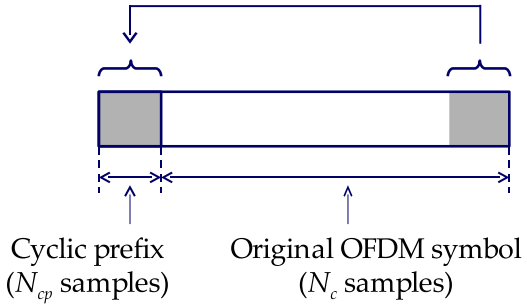


Figure 2.5: Cyclic prefix

operations [100].

The multi-carrier modulation is essentially equivalent to applying the *inverse discrete Fourier transform* (IDFT) on the original high rate data stream. Similarly, the multi-carrier demodulation at the OFDM receiver can be achieved by a *discrete Fourier transform* (DFT). The DFT functionality can be implemented via computationally efficient *fast Fourier transform* (FFT) algorithms [14]. Recent advances in the semiconductor industry enabled cost effective implementations of large size FFT chips, paving way for the deployment of commercial OFDM systems.

2.3.6 Cyclic Prefix

In a flat fading environment, the orthogonality between the subcarriers is maintained and the transmitted signals can be reconstructed perfectly at the receiver. However, when the OFDM signal is transmitted over a multipath fading channel, the time dispersion of the channel leads to the loss of orthogonality between the subcarriers, and ICI and ISI will be introduced [83].

For the purpose of eliminating the ISI, an empty guard interval could be introduced. As long as this guard interval is longer than the maximum delay spread of the channel, the orthogonality between adjacent symbols on the same subcarrier will be preserved. However, the introduction of an empty guard interval will not help eliminate the loss of orthogonality between subcarriers caused by the multipath channel.

This problem can be overcome by using a cyclic prefix [72]. The cyclic prefix is constructed by copying the last part of the OFDM symbol and prefixing it as

guard interval at the beginning of the OFDM symbol. Recall that an OFDM symbol consists of N_c samples (i.e., the output of the IFFT). The cyclic prefix is made up of last few, say N_{cp}, samples of the original OFDM symbol (see Figure 2.5).

Due to the cyclic prefix, the OFDM symbol is cyclically extended into the guard interval. The transmitted signal appears periodic and the effect of the time dispersion becomes equivalent to a cyclic convolution. The FFT interval can be selected to ensure that each subcarrier has an integral number of cycles during the FFT interval, if the cyclic prefix is longer than the maximum multipath delay spread. The orthogonality of the subcarriers is maintained, eliminating the ICI and ISI caused by the time dispersion of the channel.

Since the cyclic prefix is an extension of the OFDM symbol, it does not carry any new information. This reduces the efficiency of the system by a factor of N_{cp} / $(Nc + N_{cp})$ [39]. For an OFDM system employing a high number of subcarriers, this loss will not be significant. Furthermore, the redundancy information contained in the cyclic prefix can be exploited in channel estimation and synchronization.

2.3.7 Channel Estimation

The estimates of the channel state information are used at the receiver for compensating the amplitude and phase variations on the received symbols. When high-order modulation schemes such as M-ary QAM [78] are used, the accuracy of the channel state information is crucial in detecting the symbol constellation points accurately.

The channel state information for each subcarrier can be estimated with the use of pilot/training symbols [31, 25, 86]. Known symbols are inserted into the transmitted signal and the receiver exploits the knowledge of these known symbols. Pilot based estimation methods incur loss in utilization due to the overhead involved [95]. Blind channel estimation methods estimate the channel state information without the help of the pilot symbols. They either use the existing redundancies in the system such as the cyclic prefix [41, 91, 61] or impose certain struc-

tures on the transmitted blocks which can be exploited by the receiver [82, 73].

2.3.8 Time and Frequency Synchronization

OFDM systems are more susceptible to time and frequency offset errors than single carrier systems [75, 63]. The time offset error is caused by the incorrect identification of the OFDM symbol boundary at the receiver introducing ISI and ICI. The frequency offset is caused by the mismatch of the carrier frequencies in the transmitter and the receiver and exacerbated by the Doppler effect. The frequency offset error can destroy the orthogonality of the subcarriers introducing ICI.

The inclusion of the cyclic prefix makes OFDM relatively more robust to time offset errors. The time offset may vary over an interval equal to the length of the cyclic prefix without causing ICI or ISI. On the other hand, OFDM is relatively more sensitive to frequency offset errors [76].

The objective of time synchronization is to estimate where the symbol boundary lies, so that an uncorrupted portion of the received OFDM symbol can be sampled for FFT. Frequency synchronization refers to the process of estimating the frequency offset of each subcarrier and compensating for it to regain subcarrier orthogonality.

Various approaches have been proposed in the literature for time and frequency synchronization in OFDM systems. These techniques can be broadly classified as either pilot based or blind. The pilot based synchronization methods [110, 54, 26] require pilot symbols to be included in the OFDM symbol while the blind synchronization methods [97, 18, 62] essentially exploit the redundancy information contained within the cyclic prefix. The pilot symbols and cyclic prefixes, like the data symbols, are impaired by the time dispersion and the frequency distortion of the channel. The synchronization information are extracted by performing cross-correlation calculations on cyclic-prefixes or auto-correlation calculations on the pilot symbols.

2.3.9 Peak to Average Power Ratio Reduction

High *peak to average power ratio* (PAPR) has been recognized as one of the major problems involving OFDM modulation. Because the OFDM signal is formed by adding a large number of independent subcarrier tones, the peak values of some transmitted signals could be much larger than the typical values. To be able to handle these occasional large peaks, a high-power amplifier with linear characteristics over a large dynamic range is required. As the power consumption of a high-power amplifier largely depends on its peak power output rather than the average power output, this leads to a low power efficiency [48]. On the other hand, if the high power amplifier does not handle high peaks (i.e., behaves nonlinearly), the resulting signal clipping will cause out-of-band radiation and degradation of the performance [71].

Various techniques have been proposed for reduction of PAPR. The simplest method to reduce the PAPR is to perform explicit amplitude clipping (setting peak amplitudes to a predetermined level) and filtering to suppress the out-of-band radiation [67, 57]. However, this technique suffers from various problems such as peak regrowth, and requires additional processing [64, 9]. Another class of PAPR reduction methods uses coding techniques [30, 70]. The idea is to avoid the code words (bit combinations) that give rise to OFDM signals with high PAPR. Other PAPR reduction techniques include tone reservation [49, 7] and selective mapping [11, 20].

Chapter 3

Resource Allocation in OFDM Networks

3.1 Introduction

Resource allocation schemes for multiuser OFDM systems can benefit by taking account of various forms of diversity such as frequency diversity and multiuser diversity. Frequency diversity refers to different subcarriers within a wireless link having different channel gains due to the frequency selective nature of the channel, while multiuser diversity refers to different users experiencing different channel conditions due to their different locations in the network. These diversities imply that a subcarrier that is in a deep fade for one user may not be in deep fade for the other users. By allocating the subcarriers to users based on the channel conditions the users see on the subcarriers, these diversities can be exploited.

Wahlqvist *et al.* [99] showed that dynamic (adaptive) resource allocation can improve the performance of OFDM systems. Adaptive resource allocation strategies allow available resources to be used efficiently, by taking account of the channel conditions, and allocate, to each user, a subset of subcarriers which experience good channel conditions for that user. The users can transmit simultaneously, in parallel. In this context, *Orthogonal Frequency Division Multiple Access* (OFDMA) simplifies the allocation problem. OFDMA is defined as one in which users are assigned different subsets of the subcarriers so that the users are mutually orthogonal.

In this chapter, we contextualize our dynamic research allocation problem by surveying various adaptive resource allocation schemes in the literature. The outline of the chapter is as follows. Sections 3.2 - 3.4 review the adaptive resource allocation schemes that adapt the resource allocation to instantaneous channel

conditions. Section 3.2 describes the classical water-pouring technique for the single user systems. OFDMA based multiple user resource allocation techniques in the context of single cell systems (i.e., no inter-cell interference) are presented in Section 3.3. Section 3.4 reviews the existing work on multiple cell resource allocation in the literature. The use of knowledge of the statistics of the channel in the resource allocation as opposed to using perfect knowledge of instantaneous channel conditions is justified in Section 3.5. Section 3.6 defines our partial channel state information based resource allocation problem and describes our two layer approach to solve it. Fast frequency hopping based interference averaging which we make use of in our design, is presented in Section 3.7.

3.2 Single User Resource Allocation

In a single user OFDM system, the user has full access to all subcarriers. In this context, the resource allocation problem is to determine an appropriate transmit power allocation across the subcarriers. When the perfect channel state information (CSI) is available at the transmitter, the power allocation for the user across orthogonal subcarriers with additive white Gaussian noise can be obtained with the *water pouring* [27] method. Water pouring finds the optimal power allocation across the subcarriers that maximizes the capacity for a given total power constraint.

When the power allocation is determined by the water pouring technique, power P_i allocated to subcarrier i satisfies $P_i + \sigma_i^2 = B$ where σ_i^2 is the "effective" noise spectral density on subcarrier i (i.e., noise spectral density divided by the fading coefficient of subcarrier i) and the "water level" B is a constant that is the same for all subcarriers. The more the effective noise spectral density on a subcarrier is, the less the power allocated to that subcarrier will be. Since P_i is a non-negative quantity, if $\sigma_i^2 \geq B$, then $P_i = 0$, i.e., subcarrier i will not be used for transmission at all. The value of B is determined by solving the equation $\sum_{i=1}^{N_c} (B - \sigma_i^2)^+ = P_{total}$, where N_c is the number of subcarriers in the system, $x^+ = \max(x, 0)$ and P_{total} is the power budget.

3.3 Single Cell, Multi-user Resource Allocation

In the context of multiple user OFDMA networks, the resource allocation involves assignment of subcarriers to users and selecting the transmit powers to be used on the allocated subcarriers. It is possible to perform the multiuser resource allocation by using water pouring based techniques. Yu *et al.* [107] proposed an iterative water pouring algorithm to solve a problem of maximizing the sum capacity for a Gaussian multiple access channel. Jang and Lee [43] solved a data rate maximization problem for a multiuser OFDM network, by assigning each subcarrier to its best user (i.e., the user with the highest channel gain), and then running water pouring among the subcarriers assigned to each user individually. However, these water pouring based approaches do not, in general, guarantee fairness among users. The users that are further away from the base station or with a bad channel quality will be disadvantaged.

In order to support the individual rate or quality of service (QoS) requirements, these requirements need to be included into the resource allocation optimization problem. Such resource allocation problems in the literature can be broadly classified into two categories, based on the optimization objectives. The objective in the first category of work is to minimize overall transmit power given individual user quality of service (QoS) requirements (e.g., data rate requirements). The body of work in the second category aims to maximize the throughput under a power constraint with some fairness criteria (e.g., proportional fairness among the users). Both categories of problems are combinatorial in nature due to the discrete nature of the subcarrier allocation.

3.3.1 Power Minimization Problem

Consider an OFDMA system with M users and N_c subcarriers. User m has a data rate requirement of R_m^{tar}. Let $\Gamma_m(i)$ be the channel gain, $p_m(i)$ be the transmission power, $\sigma_m^2(i)$ be noise spectral density for user m on subcarrier i. Then, the signal-to-noise ratio (SNR) on subcarrier i at receiver m is $\gamma_m(i) = \Gamma_m(i)p_m(i) \;/\; \sigma_m^2(i)$. The transmission rate $r_m(i)$ of subcarrier i for user m is a function of the SNR, that

is $r_m(i) = f_m(\gamma_m(i))$. Function $f_m(.)$ is monotonically increasing and $f_m(0) = 0$. Note that if subcarrier i is not allocated to user m then $p_m(i) = 0$ and $r_m(i) = 0$.

The power minimization problem is to determine a subcarrier allocation to users and a power allocation for the subcarriers that minimize the total transmit power while satisfying the user rate requirements:

$$\min \sum_{m=1}^{M} \sum_{i=0}^{N_c} p_m(i) \tag{3.1}$$

subject to

$$\begin{aligned} \sum_{i=1}^{N_c} r_m(i) &\geq R_m^{tar}, \quad m = 1, 2, \ldots, M \\ \sum_{m=1}^{M} \delta\left[r_m(i)\right] &\leq 1, \qquad i = 1, 2, \ldots, N_c, \end{aligned}$$

where $\delta[.]$ is the Kronecker delta function.

The optimization within (3.1) is a non-convex problem and combinatorial in nature, making it computationally intractable. Hence, various suboptimal approaches have been proposed to solve it in the literature. Wong *et al.* [101] apply a Lagrangian Relaxation (LR) to this problem. They relax the problem to allow time sharing of subcarriers between users. With this relaxation, the problem can be turned into a convex problem. This relaxed problem is solved with an iterative search algorithm based on Lagrangian techniques. Once the solution to the relaxed problem is found, discrete subcarrier allocation is obtained by assigning each subcarrier to the user who has the largest time share on that subcarrier. Then, using the assigned subcarriers, single user bit loading (power allocation) is applied to each user. Firstly, the iterative search algorithm takes a large number of iterations to converge, and the accuracy and the rate of convergence are highly dependent on the choice of the step size. Secondly, some users may be disadvantaged with the quantization method used. A simplified version of the algorithm was also proposed by Wong *et al.* [102], in which the modulation scheme for each user is fixed across all subcarriers.

An alternate way to reduce the complexity of Problem (3.1) is to decompose it using a layered approach. A two step algorithm was proposed by Kivanc *et al.* [47, 46] to solve Problem (3.1). The first step of the algorithm is the bandwidth allocation, which determines the number of subchannels to be allocated to each user based on the average SNR of the users. The second step then allocates the required number of subcarriers (determined by the first step) to users, by applying a greedy approach considering the channel quality of the individual subcarriers. In this step, the algorithm examines the subcarriers, assigning one subcarrier at a time to the user who has the highest SNR on it and has not yet received the required number of subcarriers. The subcarrier allocation produced by this algorithm is not unique, and depends on the order in which the subcarriers are considered. Reference [58] refines the second step of this algorithm to alleviate the above mentioned problem at the expense of increased complexity.

Kim *et al.* [44] also use a layered approach to solve the resource allocation problem. Firstly, the subcarrier allocation is determined by assuming the same modulation scheme for each user across all subcarriers. This involves solving a linear programming problem. Once the subcarrier allocation is determined, the bit loading on the subcarriers is done using a greedy algorithm.

References [52, 89] consider variations of Problem (3.1), in which a sum-rate requirement is imposed instead of individual rate requirements.

3.3.2 Rate Maximization Problem

The problem of rate maximization with fairness was also widely studied in the literature. Rhee and Cioffi [81] considered a rate maximization problem with the objective of maximizing the minimum rate among the users, for a given power budget. By allowing continuous allocation of subcarriers, it was shown that the problem is convex and the optimal assignment can be found using numerical methods. The proposed low-complexity adaptive subchannel allocation uses a flat power spectral density (PSD) across all subcarriers. Subcarriers are allocated to users, one at a time, with the worst user (i.e., with the lowest achieved rate)

getting the subcarrier with the highest gain. Armada [8] studied a problem of maximizing the number of users with equal rate requirements, for a given total power. Manoharan and Bhashyam [60] and Shen *et al.* [85] extend the rate maximization problem to include proportional fairness.

3.4 Multi-Cell, Multi-User Resource Allocation

In a single cell OFDM network, the subcarrier allocation can be performed in an OFDMA fashion (i.e., each user can be assigned an exclusive subset of subcarriers) as the number of users is small compared to the number of subcarriers available in the system. It may not be possible to extend this idea to a multiple cell OFDM system in order to allocate an exclusive subset of subcarriers to every user in the network. This would require a great deal of coordination amongst the cells, which is not realistic. Even if it were possible, it may not be desirable, as it does not take advantage of the principle of frequency reuse.

Some form of frequency reuse is required to handle the multicell resource allocation problem. It can either be a full-reuse or a partial-reuse. With full frequency reuse, all subcarriers are available in each cell. In this case, the users in each cell will experience the interference from every other cell in the network. On the other hand, with partial reuse, adjacent cells in the network are grouped together and each group will have full access to the subcarriers. Within each group, the subcarriers are partitioned between cells exclusively so that all users within each group are mutually orthogonal. Partial frequency reuse is similar to the frequency planning concept [53] in the traditional cellular networks. The major difference is that, unlike in the traditional cellular networks, cells in a group can be allocated non-contiguous subcarriers in OFDM networks. Regardless of whether full or partial reuse is employed, the resource allocation needs to take account of the inter-cell interference.

Various multicell resource allocation schemes for OFDM systems have been proposed in the literature, which adapt the allocations based on instantaneous channel conditions. Zhang *et al.* [109] consider a resource allocation problem for

a cellular OFDM system. The problem is solved in two stages: adaptive cell selection and intra-cell resource allocation. The adaptive cell selection involves choosing a serving cell (base station) for each user based on the average received power strengths at the receiver and the traffic load at the candidate cells. At the cell level, the resource allocation problem is formulated as rate maximization while achieving the user rate requirements for a given power budget. Equal distribution of powers to the subcarriers reduces the problem to a linear integer problem (the original problem is a nonlinear integer problem). The proposed suboptimal algorithm finds the subcarrier and bit allocation in a finite number of operations. However, the inter-cell interference is not explicitly modeled.

Pietrzyk *et al.* [74] solve a problem similar to that of [109] in three steps: cell selection, subcarrier allocation, and bit allocation. The criterion for cell selection incorporates both channel gains and interference levels of the subcarriers. The subcarrier allocation phase determines the allocation of individual subcariers to users. The number of subcarriers for each user is proportional to its rate requirement; the individual subcarriers are then allocated to users based on a greedy approach. Finally, the bit allocation is performed by taking the co-channel interference into account.

Kulkarni *et al.* [50] consider a resource allocation problem for a infrastructureless wireless network. The network consists of point-to-point wireless links (set of transmitter-receiver pairs). The allocation problem is to identify sets of links that share each subcarrier and find the bit loading on it, with the objective of minimizing the total transmit power subject to meeting the rate requirements of the links. As this problem is computationally intractable, heuristic algorithms are proposed. They propose a distributed algorithm which performs the allocation of subcarriers for each link by considering the carrier-to-interference ratio of the subcarriers. A subcarrier i^* is chosen for a link l if the rate requirement of link l has not already been met and subcarrier i^* has the highest carrier-to-interference ratio, i.e.,

$$i^* \leftarrow \max_i \frac{G_l(i)}{I_l(i)},$$

where $G_l(i)$ and $I_l(i)$ are the channel gain and the interference respectively, for link l on subcarrier i. Once a subcarrier is selected for a given link, an iterative power control is initiated in order to find an appropriate power for the newly selected subcarrier on the link of interest as well as to adjust the powers for other links that are already using this subcarrier. While the interference is explicitly modeled, the bit loading on the subcarriers are done independently and thus coding across subcarriers is not taken into account. Han *et al.* [38] take a game theoretic approach to solve a similar problem. Kwon and Lee [51] also adopt a game theoretic approach, but to solve a downlink resource allocation problem in a multi-cell OFDMA system.

Damji *et al.* [28, 29] adopt a *bandwidth-constrained power minimization (BCMP)* strategy to mitigate inter-cell interference in cellular networks. The idea is to first place an upper bound on the number of subcarriers to each user based on the rate requirements and then adjust the transmit powers accordingly.

Heo *et al.* [42] propose a reuse partitioning based resource allocation for an OFDM cellular network. Each cell consists of three sectors that are also divided into inner and outer cell. The frequency reuse factor (FRF) for the cell is one. The available subcarriers are grouped into two pools: subcarriers with FRF 1 and FRF 3. The subcarriers with FRF 1 will be used in the inner cell. One-third of the FRF 3 subcarriers will be exclusively used in each sector of the outer cell.

Li and Liu [56] consider a resource allocation problem in which the base stations cooperate through the radio network controller (RNC). A hierarchical resource allocation scheme is proposed. The initial allocation is performed by the RNC with limited feedback from all base stations. This allocation is then refined locally by the respective base stations.

3.5 Availability of Channel State Information (CSI)

The resource allocation schemes that we have seen in Sections 3.2 - 3.4 perform the allocations based on instantaneous channel conditions and work with the assumption of perfect channel state information (CSI) at the transmitters. As such,

these schemes need to adapt to the instantaneous channel variations.

In a single cell OFDMA system with time division duplex (TDD), the transmitters can estimate the instantaneous channel conditions due to the reciprocity of the uplink and the downlink channels. On the other hand, in an OFDMA system with frequency division multiplex (FDD), there is no channel reciprocity since the uplink and the downlink fade independently. The channel state information needs to be conveyed by the receiver to the transmitter.

In a multicell OFDM system, the assumption of perfect CSI at the transmitters is not realistic regardless of whether the system uses a TDD or FDD mode. The interference level at the receiver of each subcarrier will need be fed back to the base station in order to have perfect transmitter CSI. Firstly, the signaling overhead due to the feedback of CSI from the receiver to the transmitter will be significant. In addition, the requirement of perfect channel state information (CSI) is only suitable for static or very slow time-varying channels such as in a wireline systems. Since mobile radio channels vary quickly in time, it may not be possible for the transmitter to keep the channel state information up to date.

The effectiveness of the dynamic resource allocation depends on the accuracy of the CSI at the base station. In cases where the feedback channel is not sufficiently reliable or the channel is changing so quickly that the transmitter cannot adequately track the channel variations, the allocation based on the channel feedback may not be effective.

When the perfect CSI is not available at the transmitter, the use of partial CSI such as long term statistics of the channel will still lead to better performance compared to fixed allocation (that does not take account of channel state information at all). The resource allocation with imperfect or partial CSI at the transmitter has not been fully explored in the context of OFDM systems.

Leke and Cioffi [57] studied the effect of imperfect channel knowledge on the performance of the single user multi-carrier systems. Two sources of error related to channel estimation were considered: error in the channel estimate and the variations in the channel after the estimation is performed. Song *et al.* [87] consider adaptive modulation schemes for OFDM systems using the fading statistics of Ri-

cian fading channels. Yao and Giannakis [105] perform the capacity analysis of an OFDM system with partial CSI. The deterministic mean of the channel gains and the distribution of the error component for the subcarriers are assumed known at the transmitter. Both ergodic and outage capacity are evaluated.

In [59], the authors consider the problem of power loading for the subcarriers in an OFDM/FDD system with limited feedback. With the proposed scheme, instead of feeding the CSI back to the transmitter, the receiver chooses an appropriate power loading vector and conveys this information to the transmitter.

The works above consider a single user in isolation and do not take the interference into account. The focus of this thesis, on the other hand, is to consider a resource allocation problem in a multicell OFDM cellular system in which only partial CSI is available at the transmitters (base stations). We explicitly model the inter-cell interference.

3.6 The Problem Addressed in this Thesis

This thesis investigates the resource allocation problem in a multiple cell, multiple user OFDM cellular network, and we focus on the downlink in which the base station is transmitting to mobiles. There are several mobiles in each cell. The resource allocation problem is to decide the allocation of subcarriers to each user within a cell, and the power allocation across the subcarriers in order to satisfy the data rate constraints of the users.

If the base stations had perfect channel state information (CSI) of all the users in the cellular system (i.e., perfect knowledge of the instantaneous channel conditions of all the users in the entire network and the transmit powers of other base stations) this knowledge could be used in solving the above resource allocation problem. As we have seen in Sections 3.3 and 3.4, this problem is fundamentally combinatorial in nature (even in the context of the single cell networks), and this leads to a search for suboptimal solutions that are computationally efficient. In a multiple cell network, there is the additional problem of how to distribute the calculations, given that cell sites will almost certainly not know the channel

conditions of users outside their cells.

The assumption of perfect CSI is not realistic, when the channel conditions change quickly with time, and more so in the presence of the inter-cell interference. As we have discussed in Section 3.5, in order for the base station to have perfect CSI of all subcarriers for each of its users, the amount of feedback information required will be significant. Furthermore, it may even not be possible to track the instantaneous channel conditions at the base stations in real time when the channel changes quickly in time. The resource allocation based on the feedback on the instantaneous channel conditions may have become out of date due to the changes in the channel condition since the last measurement was taken. Additionally, in a multi-cell environment, it is not sufficient to just measure the channel conditions and report back to the base station as in the case of a single cell problem with known channel state. Since users in other cells will also be updating their transmit powers and subcarrier allocations, the interference cannot be treated as a constant.

When perfect CSI is not available, the resource allocation needs to be performed based on knowledge of the statistics of the channel conditions. In this case, it is not possible to perform resource allocation to ensure that users receive the target data rates at all times; if a user is unlucky and its allocated subcarriers fade simultaneously, it will not be able to transmit reliably at its target data rate and the user will be in outage. The outage probability is then the appropriate metric to measure the performance of the users.

One particular problem for the design of any outage probability based resource allocation algorithm is the difficulty of characterizing the outage probability as a function of all the system parameters (transmit powers on each subcarrier, number of subcarriers allocated to each user, interference experienced on each subcarrier, *etc.*). Since there is interference between users in different cells, the resource allocation algorithm will need to take this interference into account, and such interference will itself be changing as the algorithm converges to equilibrium. The algorithm needs to learn statistics from direct measurements in order to track the inter-cell interference from the other cells, in addition to be able to

track the changes within the cell (user mobility, call arrivals and departures *etc.*). Solving this problem exactly appears to be intractable.

We solve the outage based resource allocation problem by taking a layered approach, and make use of the fast frequency hopping based *interference averaging* technique in our design. We adopt a two timescale, layered approach. On a slow timescale (higher layer), per-link transmit powers are updated by taking into account inter-cell interference. The power updates are based on feedback from the receivers. The feedback is computed based on measured average channel conditions, averaging over the frequency selective fading. On a fast timescale (lower layer), subcarriers are allocated to each link taking into account the statistics of the fast fading and the different number of subcarriers each user can average over.

In the presence of inter-cell interference, the system design should avoid the possibility of some users being adversely affected by dominant interferers. For this purpose, we employ the fast frequency hopping based interference averaging technique, which will be introduced in Section 3.7. With interference averaging, the interference for a user comes from all users in the neighbouring cells instead of a particular subset of users.

The subcarrier allocation within each cell is done in a mutually exclusive fashion. Therefore, the users in the same cell do not interfere with each other. However, all subcarriers are reused in every cell, i.e., a frequency reuse factor of one is used. Consequently, there is "cross-talk" between users in different cells due to inter-cell interference.

3.6.1 Higher Layer Resource Allocation

The higher-layer resource allocation problem takes account of the coupling that exists between the cells due to the interference created by the inter-cell users. In other words, this layer accounts for the fact that if one user increases its transmit power on any subcarrier, then this will increase the interference experienced by all other users that experience this user as an interferer. In this step, we ignore the fading, and solve an initial deterministic resource allocation problem that ac-

counts only for the average gains on each link.

The higher-layer resource allocation problem is greatly simplified by taking advantage of the method of interference averaging. This allows us to model the allocation of individual transmit powers to each link instead of each subcarrier ("receiver-directed power control", as opposed to "subcarrier-directed" power control). This corresponds to a resource allocation problem in a flat fading environment. The subcarrier allocation to users is characterized by a proportion of subcarriers ("bandwidth") allocated to each user, and the power spectral density (PSD) for that user, without the need to model the exact choice of subcarriers.

Chapters 4 and 5 of this thesis are devoted to the higher-layer deterministic resource allocation problem. The optimal scheme (Chapter 4) allocates a different power spectral density to each link (receiver-directed power control). We also investigate the impact of reducing the complexity of this problem by adding additional constraints. A "static subcarrier allocation" scheme (Section 5.2) and a "uniform transmit PSD allocation" scheme (Section 5.3) are considered in order to reduce the complexity. The latter scheme reduces the dimension of the problem so that each cell has a single scalar power variable; power is allocated uniformly across all subcarriers within the cell, rather than on a per link basis. It is shown that in the deterministic model, there is very little loss from this approach.

The "uniform PSD" power control policy provides the appropriate structure to apply the lower layer algorithm to handle frequency selective fading.

3.6.2 Lower Layer Resource Allocation

The reason for the fast time-scale algorithm is the following: when the deterministic resource allocation is applied directly to a time-varying, frequency selective fading environment, the resulting resource allocation may not lead to satisfactory performances of the individual users due to:

a) some users having particularly high outage probabilities and/or

b) an uneven distribution of outage probabilities among the users.

Adding a rate margin (i.e., targeting higher data rates for the users) during the running of the higher layer resource allocation can improve the outage probabilities of the users. Note that rate margins have a limitation: there is a knock-on effect from the interference, i.e., adding a rate margin for a user in a particular cell will increase the interference for users in other cells, which needs to be combated by adding appropriate rate margins for those affected users as well.

We can also add a margin by allocating more subcarriers to a user. This corresponds to redistributing the subcarriers among the users within the cell (since the total number of subcarriers is fixed) which is the basic idea behind the lower layer approach. In particular, we focus on the lower layer problem of finding a subcarrier allocation that "balances" the outage probabilities (i.e., minimizes the maximum outage probability) among the users by making use of the statistics of the channel conditions.

Once the transmit powers are found by the high layer optimization, then further optimization of subcarrier allocations for a given user can be performed without the need to keep track of interference updates from other cell users. If the transmit power allocations are fixed, then the other interferers can update their subcarrier allocations as they wish, without affecting the given user. Furthermore, outage probability calculations can be based on a central limit theorem approximation. The use of interference averaging is central to this approximation: it removes the need to associate each subcarrier with a particular interferer, and ensures that "all subcarriers are equal".

3.7 Interference Averaging

When a user is allocated a particular subset of subcarriers, the interference experienced by that user comes from just the signals in these subcarriers from the nearby cells. When the interference on these particular subcarriers is strong, the user will be severely affected. In a system with full or aggressive frequency reuse, it may not be possible to find an allocation of subcarriers that avoids such a possibility. Therefore, the best possible strategy to reduce this impact is to ensure

that the interference for a user is not attributable to a single user, but is an average from many users. This is the basic idea behind the notion of interference averaging.

The spread spectrum systems achieve interference averaging by making the signal occupy the entire system bandwidth either through spreading or hopping. The interference averaging is an inherent feature of CDMA systems, due to the use of spreading codes to spread users' signals over the entire frequency spectrum. In OFDM systems, the interference averaging can be achieved by making use of predetermined fast frequency hopping patterns.

In fast frequency hopping OFDM systems, a fast hopping pattern across the subcarriers will define a virtual subchannel. The number of virtual subchannels in a cell is equal to the number of subcarriers in the system. The orthogonality between virtual subchannels within a cell can be retained by implementing Latin Square design based fast frequency hopping.

3.7.1 Latin Square Design Based Fast Frequency Hopping

A *Latin square of order* N_c is an $N_c \times N_c$ matrix with entries from a set S of N_c distinct elements, say $S = \{0, 1, \ldots, N_c - 1\}$ such that each of the N_c elements of S occurs once (and hence exactly once) in each row and once in each column.

Let N_c represent the number of subcarriers in the system. Suppose that the hopping pattern has a period of N_c symbol periods (i.e., hopping pattern is repeated after every N_c symbol periods). This periodic hopping pattern across the subcarriers in a cell can be represented by a $N_c \times N_c$ matrix with entries from the set of N_c virtual subchannels. Each row of the hopping matrix corresponds to a subcarrier and the entries represent the virtual subchannels that occupy that subcarrier at different symbol periods. For example, the (i, j) entry of the matrix corresponds to the virtual subchannel which is occupying the i^{th} subcarrier at j^{th} symbol period. A hopping matrix needs to meet the following design requirements:

1. Intra-cell orthogonality: At any symbol period, the virtual subchannels oc-

cupy different subcarriers.

2. Maximal frequency diversity: Each virtual subchannel hops over all subcarriers during a hopping cycle.

These two requirements together correspond to the joint constraint that each row and each column of the hopping matrix contains each virtual subchannel exactly once. A Latin square of order N_c will meet these requirements.

Figure 3.1 shows examples of Latin square design based hopping patterns for $N_c = 7$. The corresponding hopping matrix is shown below:

$$\begin{bmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 3 & 4 & 5 & 6 & 0 & 1 \\ 4 & 5 & 6 & 0 & 1 & 2 & 3 \\ 6 & 0 & 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 3 & 4 & 5 & 6 & 0 \\ 3 & 4 & 5 & 6 & 0 & 1 & 2 \\ 5 & 6 & 0 & 1 & 2 & 3 & 4 \end{bmatrix}.$$

Note that the subcarrier and symbol period indices run from 0 to 6 (and not from 1 to 7). Each virtual subchannel will go through all physical subcarriers during a hopping cycle. For example, the virtual subchannel 5 will occupy subcarrier 6 in the first symbol period (number 5 appears at the (6,0) entry), 2 in the second, 5 in the third, and so on.

In order to achieve the interference averaging, we require that, during one hopping cycle, any two virtual subchannels in adjacent cells have minimal overlap, i.e., use the same subcarrier in the same symbol period exactly once. This corresponds to the hopping matrices of the neighbouring cells being *orthogonal*. Two Latin squares A and B are said to be orthogonal if the ordered pairs obtained from the juxtaposed array $A \times B$ exhaust all N_c^2 possibilities, i.e., every ordered pair occurs exactly once.

It is possible to construct a set of $N_c - 1$ mutually orthogonal hopping matrices of order N_c, if N_c is a prime or a power of a prime [21]. If N_c is a prime, the

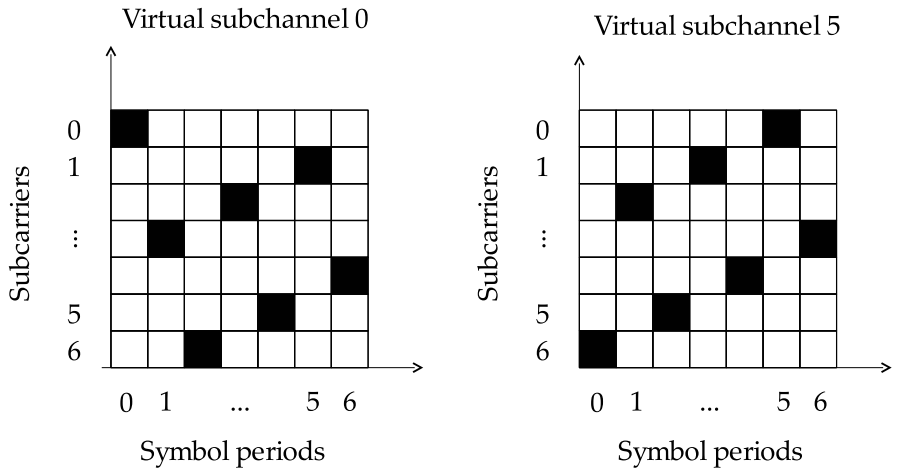


Figure 3.1: Virtual subchannel hopping patterns for $N_c = 7$. The hopping patterns for virtual subchannels 0 and 5 are illustrated.

construction of these mutually orthogonal hopping matrices can be done in a simple manner. For $a = 1, \ldots, N_c - 1$, we define an $N_c \times N_c$ matrix L^a with (i, j) entry

$$L^a(i, j) = ai + j \text{ (modulo } N_c \text{).}$$

Note here that $i = 0, \ldots, N_c - 1$ and $j = 0, \ldots, N_c - 1$.

With Latin square based design, the inter-cell interference seen by a virtual subchannel, comes from all virtual subchannels (or equivalently, from all subcarriers) in each hopping cycle. In order to fully utilize the interference averaging, the users within the cell must be synchronized to their corresponding base station at the OFDM symbol level. Furthermore, each base station must also be synchronized.

The frequency hopping based on Latin squares was originally proposed by Wyner [104] and Pottie and Calderbank [77]. Flarion's FLASH-OFDM [2] implements this frequency hopping to provide frequency diversity, and to average the inter-cell interference, for both the uplink and downlink. The frequency re-use factor is 1, so low rate codes are used to combat relatively strong levels of intercell interference. Mobiles within a cell are allocated one or more channels (depending on their data rate requirement) and thus users are orthogonal within a cell.

Chapter 4

Resource Allocation Under Flat Fading

4.1 Introduction

As we outlined in Section 3.6, the focus of this thesis is to solve a downlink resource allocation problem in an OFDM cellular network in a multipath frequency selective fading environment. A two-layer approach is proposed to solve this problem. The higher layer resource allocation is responsible for obtaining an initial allocation by tracking the slow timescale channel changes such as distance based attenuation and log-normal shadowing. The lower layer allocation then refines the resource allocation from the higher layer, by taking account of the fast fluctuations due to frequency selective fading.

The higher layer resource allocation is performed based on the feedback from the users on the measured average channel gains. The higher layer resource allocation problem is equivalent to a problem in a flat fading environment, which is the topic of this chapter and Chapter 5. While this chapter solves the flat fading environment allocation problem exactly, the reduction in complexity of the allocation is the topic of Chapter 5. The lower layer resource allocation which refines the result of higher layer resource allocation with the aid of the statistics of the channel variations, is covered in Chapter 6.

The specific optimization problem associated with the flat fading environment is to determine the subcarrier allocation to users and the power allocation across the subcarriers, that minimizes the aggregate transmit power of the base stations, subject to meeting the data rate requirements of all users. The rationale behind this particular formulation is not to particularly conserve transmit power at the base stations. Rather, by minimizing power, we also maximize the spectral effi-

ciency of the network, as it is an interference limited system. Thus, our primary concern is spectral efficiency.

We take advantage of the method of *interference averaging* which greatly simplifies our resource allocation problem. As described in Section 3.7, interference averaging in OFDM systems can be achieved by using fast frequency hopping which effectively spreads the signal over the entire available bandwidth. Interference averaging avoids the possibility of some users being adversely affected by dominant interferers, by explicitly whitening the interference seen by any user. Each user experiences an average level of interference determined by *all* the signals from all the users in the nearby cells, not just other-cell users using particular subcarriers allocated to the user of interest. As a consequence of that, the resource allocation to users can be modeled as a proportion of subcarriers allocated to each user, and the transmit power spectral density (PSD) for that user, instead of modeling the exact choice of subcarriers. The bit loading across the allocated subchannels for any particular user will be the same, allowing coding across subchannels.

The outline of this chapter is as follows. Section 4.2 develops the system model for the higher layer (i.e., flat fading environment) resource allocation problem, which will be used in this chapter as well as in Chapter 5. This model treats proportions of subcarriers to users as continuous variables. The resource allocation problem is formulated as a power minimization problem in Section 4.3. In Section 4.4, the resource allocation problem is transformed into a convex optimization problem by the introduction of an appropriate change of variables. We show that solving the transformed problem using convex optimization tools does not lead to a distributed solution as this approach requires explicit cooperation between cells. The focus of Section 4.5 then is to design a distributed algorithm to solve the resource allocation problem without requiring explicit cooperation between cells. An iterative algorithm is devised and its convergence proven. In Section 4.6, we consider the discrete version of the resource allocation problem in which the proportions of subcarriers to users are discrete. A distributed algorithm is developed using a strategy similar to that of Section 4.5.

4.2 Model

In this section, we develop the system model and configurations which will be used in this chapter and Chapter 5. We focus on the downlink transmission under a flat fading assumption.

Time and frequency synchronization in OFDM receivers are essential for proper operation. Additionally, time synchronization at the symbol level between base stations is required for systems that employ fast frequency hopping. As synchronization issues are out of the scope of this thesis, we assume synchronization exists between the mobile users and their corresponding base stations, and also between the base stations. We note that maintaining such synchronization is feasible, e.g., Flarion's FLASH-OFDM [2].

In the resource allocation problem that we consider, the resources are the bandwidth (subcarriers) and transmit power levels, which provide degrees of freedom for optimization.

We consider the systems that use interference averaging. We note that in a flat fading environment, interference averaging can be achieved either by explicit averaging, such as the Latin square design based fast frequency hopping (as discussed in Section 3.7), or by implicit averaging, i.e., by constraining the transmit power on each subcarrier to be equal. Although, both result in an equal average interference, the instantaneous interference varies in the first case, but remains constant in the second.

We use the term *subchannel* to refer to a virtual subchannel in a system that employs explicit interference averaging, or a physical subcarrier in a system which uses implicit interference averaging. When interference averaging is employed in a flat fading environment, the average interference experienced by each subchannel is equal.

Within a cell, subchannels are allocated to users in a mutually exclusive fashion. In a system with implicit interference averaging, this corresponds to exclusive allocation of subcarriers to users. On the other hand, in a system that employs explicit interference averaging, each subcarrier is exclusively assigned to a

single user at any given time instant (as determined by the fast frequency hopping patterns). In both scenarios, the users in the same cell do not interfere with each other. The inter-cell interference arises from the "cross-talk" between users in different cells. We assume that there are several users in each cell. This may model a wideband cellular OFDM system with full or aggressive frequency reuse so that co-channel interference is significant.

As a result of employing interference averaging, we only need to model the proportion of subchannels ("bandwidth") allocated to each user and the transmit power spectral density (PSD) for that user, but not the exact choice of subchannels. The total power allocated to a user is the product of its transmit PSD and its bandwidth.

Consider a cellular network which consists of a set of N base stations, denoted by $\mathcal{N} = \{1, 2, \ldots, N\}$, with each base station $n \in \mathcal{N}$ having a set C_n of users. The apportionment of subchannels in cell n is modelled by a positive weight vector $\boldsymbol{w}_n \in \mathbb{R}_+^{L_n}$ where $L_n = |C_n|$ is the number of users in cell n. Each base station is assumed to have the same total number of subcarriers (and consequently the same number of subchannels) available. A base station n allocates a proportion $w_{n,m}$ of subchannels to its user m. This places the feasibility constraint on the weight vector:

$$\sum_{m \in C_n} w_{n,m} \leq 1.$$

To achieve maximum spectral efficiency, it will be used with equality:

$$\sum_{m \in C_n} w_{n,m} = 1. \tag{4.1}$$

Assume that the number of subchannels is large compared to the number of users per cell, allowing a continuous allocation of these resources (the handling of discrete weights is addressed in Sections 4.6 and 5.7). Let $\boldsymbol{w}$ denote the N-tuple of such weight vectors for the network. Normalized bandwidths are used; consequently, the weights are both the proportions of bandwidth and the amounts of bandwidth allocated to users.

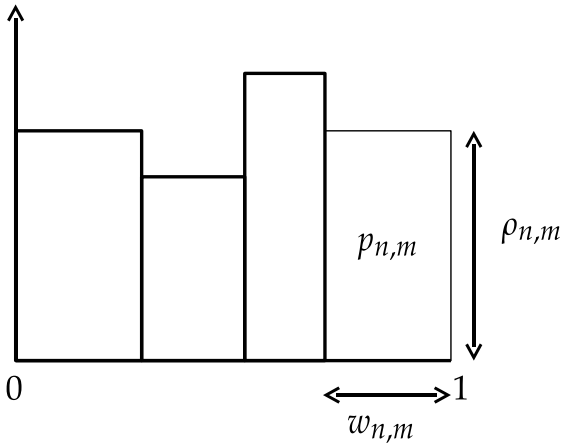


Figure 4.1: The power allocation $p_{n,m}$ to user $m \in C_n$ depends on power density $\rho_{n,m}$ and the proportion $w_{n,m}$ of bandwidth (subchannels) allocated to that user, $p_{n,m} = w_{n,m}\rho_{n,m}$. Pictorially, it is the area of the rectangle corresponding to user m. The total power q_n at base station n is the total area under the curve. With a flat power spectrum, $q_n = \rho_{n,m}, \forall m \in C_n$ since $\sum_{m \in C_n} w_{n,m} = 1$.

Denote the transmit power spectral density that a base station n allocates to user $m \in C_n$ by $\rho_{n,m}$. Let $\boldsymbol{\rho}_n = (\rho_{n,m})_{m \in C_n}$ be the vector of power densities in cell n. Denote by $\boldsymbol{\rho}$ the N-tuple of such vectors for the network. The power allocated to user m by its base station n is given by $p_{n,m} = w_{n,m}\rho_{n,m}$ since we measure power spectral density with respect to normalized bandwidth. Let $\boldsymbol{p}_n = (p_{n,m})_{m \in C_n}$ be the vector of power allocation in cell n and $\boldsymbol{p} = (\boldsymbol{p}_n)_{n \in \mathcal{N}}$ be the N-tuple of such vectors for the network. The total transmit power at base station n is $q_n = \sum_{m \in C_n} p_{n,m}$. Let $\boldsymbol{q} = (q_n)_{n \in \mathcal{N}}$ be a vector of total powers for the network. If base stations use a flat power spectrum across all subcarriers, then, $q_n = \rho_{n,m}, \forall m \in C_n$ since $\sum_{m \in C_n} w_{n,m} = 1$. See Figure 4.1 for an illustration.

Denote the receiver noise at user $m \in C_n$ by $\sigma_m^2 > 0$. The path gain from a base station k to a user $m \in C_n$ (not necessarily $k = n$) is $\Gamma_{k,m}$. Then, the signal to interference and noise ratio (SIR) at node $m \in C_n$ is given by

$$\gamma_{n,m}(\rho_{n,m}, \boldsymbol{q}) \quad = \quad \frac{\Gamma_{n,m}\,\rho_{n,m}}{\sigma_m^2 + \sum\limits_{n' \neq n} \Gamma_{n',m}\,q_{n'}}. \tag{4.2}$$

Note that since the bandwidth is normalized, σ_m^2 and q_n are both powers and average spectral densities, simultaneously.

The capacity of a link is determined by both the SIR and the number of subchannels available. Let $\tilde{f}(\gamma)$ be the maximum bit rate (in bits/sec) achieved given that the total spectrum available to the system W (in Hz) is allocated to the link of interest, and the SIR is γ. We assume that $\tilde{f}(\gamma)$ is a continuous, increasing function of γ, with $\tilde{f}(0) = 0$. A specific example is the function $\tilde{f}(\gamma) = W \log_2(1+\gamma)$, which applies if the link is optimal with respect to Shannon capacity. We will work with the associated spectral efficiency $f(\gamma) = \frac{1}{W}\tilde{f}(\gamma)$ (in bits/sec/Hz). Then, $f(\gamma)$ is also a continuous, increasing function of γ and $f(0) = 0$. If a user $m \in C_n$ with a bandwidth allocation of $W_{n,m}$ Hz has a rate target of $R^{tar}_{n,m}$ bits/sec, then, the rate constraint is:

$$W_{n,m}\, f(\gamma_{n,m}(\rho_{n,m}, \boldsymbol{q})) \;\; \geq \;\; R^{tar}_{n,m}. \tag{4.3}$$

Dividing (4.3) through by W, we get the normalized rate constraint:

$$w_{n,m}\, f(\gamma_{n,m}(\rho_{n,m}, \boldsymbol{q})) \;\; \geq \;\; c^{tar}_{n,m}, \tag{4.4}$$

where $w_{n,m} = W_{n,m} \,/\, W$ is the proportion of bandwidth allocated to user m and $c^{tar}_{n,m} = R^{tar}_{n,m} \,/\, W$ (in bits/sec/Hz) is the normalized rate requirement.

4.3 Problem Statement

The resource allocation we consider here is to determine optimal power and bandwidth allocation for users. We formulate this problem as a power minimization problem, subject to all users achieving their target normalized rates:

$$\min_{\boldsymbol{\rho}, \boldsymbol{w}} \sum_{n \in \mathcal{N}} \sum_{m \in C_n} w_{n,m}\, \rho_{n,m} \tag{4.5}$$

such that for all n,

$$w_{n,m}\, f\left(\gamma_{n,m}(\rho_{n,m}, \boldsymbol{q})\right) \;\; \geq \;\; c^{tar}_{n,m}, \; \forall m \in C_n,$$

$$\begin{aligned} \sum_{m \in C_n} w_{n,m} &= 1, \\ \rho_{n,m} &> 0, \ \forall\, m \in C_n, \\ w_{n,m} &> 0, \ \forall\, m \in C_n, \end{aligned}$$

where $\gamma_{n,m}(\rho_{n,m}, \boldsymbol{q})$ is the SIR at user m as defined by (4.2).

When we solve Problem (4.5) in this chapter, we use a particular f, namely, the Shannon formula, $f(\gamma) = \log(1+\gamma)$. This is done in order to make the analysis tractable, particularly when working with derivatives in showing the convexity of particular problems and in solving them using Lagrangian techniques. Note also that the scaling that results from the conversion from the logarithm to the base 2 to natural logarithm is ignored for the sake of convenience. On the other hand, when we consider the reduced complexity versions of this problem in Chapter 5, we will not assume any specific choice for the function f.

With Problem (4.5), there are $2L_n - 1$ parameters or "degrees of freedom" available for optimization in each cell: L_n transmit power spectral densities, one for each user, and $L_n - 1$ weights.

4.4 Convex Transformation

The optimization problem (4.5) is *not* convex, but can be transformed into one, by an appropriate change of variables [19], namely, with a logarithmic transformation. In this section, we perform this transformation and demonstrate that solving this transformed problem does not lead to a distributed solution and requires explicit message passing between different cells. The approach we use for this purpose derives from [68]. In Section 4.5, we develop a distributed algorithm to solve (4.5).

For convenience, we first modify Problem (4.5) by introducing a new auxiliary variable $\boldsymbol{t} = (\boldsymbol{t}_n)_{n \in \mathcal{N}}$ with $\boldsymbol{t}_n = (t_{n,m})_{m \in C_n}$. This effectively splits the rate

constraints of Problem (4.5), forming a new equivalent optimization problem:

$$\min_{\boldsymbol{\rho},\boldsymbol{w},\boldsymbol{t}} \sum_{n\in\mathcal{N}} \sum_{m\in C_n} w_{n,m}\,\rho_{n,m} \tag{4.6}$$

such that for all n,

$$\begin{aligned} w_{n,m}\,\log\left(1+t_{n,m}\right) &\geq c_{n,m}^{tar}, \quad \forall m \in C_n, \\ t_{n,m} &\leq \gamma_{n,m}(\rho_{n,m},\boldsymbol{q}), \quad \forall m \in C_n, \\ \sum_{m\in C_n} w_{n,m} &= 1, \\ \rho_{n,m} &> 0,\ \forall\, m \in C_n, \\ w_{n,m} &> 0,\ \forall\, m \in C_n \end{aligned}$$

where the second constraint is expected to be met with equality.

We now perform a transformation of coordinates on (4.6) with $\tilde{\rho}_{n,m} = \log \rho_{n,m}$, $\tilde{w}_{n,m} = \log w_{n,m}$ and $\tilde{t}_{n,m} = \log t_{n,m}$. In these new variables, the transformed problem is:

$$\min_{\tilde{\boldsymbol{\rho}},\tilde{\boldsymbol{w}},\tilde{\boldsymbol{t}}} \sum_{n\in\mathcal{N}} \sum_{m\in C_n} e^{\tilde{w}_{n,m}+\tilde{\rho}_{n,m}} \tag{4.7}$$

such that for all n,

$$\log c_{n,m}^{tar} - \tilde{w}_{n,m} - \log\log\left(1+e^{\tilde{t}_{n,m}}\right) \leq 0, \forall m \in C_n, \tag{4.8}$$

$$\sum_{m\in C_n} e^{\tilde{w}_{n,m}} - 1 = 0, \tag{4.9}$$

$$\tilde{t}_{n,m} - \log\left(\Lambda_{n,m}(\tilde{\boldsymbol{\rho}},\tilde{\boldsymbol{w}})\right) \leq 0, \forall m \in C_n, \tag{4.10}$$

where

$$\Lambda_{n,m}(\tilde{\boldsymbol{\rho}},\tilde{\boldsymbol{w}}) = \frac{\Gamma_{n,m}\, e^{\tilde{\rho}_{n,m}}}{\sigma_m^2 + \sum_{n'\neq n} \Gamma_{n',m} \left(\sum_{m'\in C_{n'}} e^{(\tilde{w}_{n',m'}+\tilde{\rho}_{n',m'})} \right)}.$$

In order to show that Problem (4.7) is convex, we need to show that the ob-

jective as well as the constraints (4.8) - (4.10) are convex. The objective function and the constraints (4.9) are clearly convex since each of them consists of a sum of convex exponentials of the variables $\tilde{\boldsymbol{\rho}}$ and $\tilde{\boldsymbol{w}}$. Showing that the constraints (4.8) and (4.10) are convex involves showing that both $h(x) = -\log\log\left(1+e^{x}\right)$ and $g_{n,m}(\tilde{\boldsymbol{\rho}},\tilde{\boldsymbol{w}}) = -\log\left(\Lambda_{n,m}(\tilde{\boldsymbol{\rho}},\tilde{\boldsymbol{w}})\right)$ are convex.

It can easily be shown that the function $h(x)$ is convex in x by using the second-order condition [19]. Writing the expression for $h'(x)$ and $h''(x)$, we obtain:

$$\begin{aligned} h'(x) &= -\frac{e^{x}}{[\log(1+e^{x})]\ (1+e^{x})} \\ h''(x) &= \frac{e^{x}\left[e^{x}-\log(1+e^{x})\right]}{[\log(1+e^{x})]^{2}\ (1+e^{x})^{2}}. \end{aligned}$$

Clearly, $h''(x) > 0$ since $e^{x} > \log(1+e^{x})$ for $x \in \mathbb{R}$ and hence $h(x)$ is convex. Turning our attention to $g_{n,m}(\tilde{\boldsymbol{\rho}},\tilde{\boldsymbol{w}})$, we can write $g_{n,m}(\tilde{\boldsymbol{\rho}},\tilde{\boldsymbol{w}})$ as:

$$g_{n,m}(\tilde{\boldsymbol{\rho}},\tilde{\boldsymbol{w}}) = \log\left(\frac{\sigma_{m}^{2}}{\Gamma_{n,m}}e^{-\tilde{\rho}_{n,m}} + \sum_{n'\neq n}\frac{\Gamma_{n',m}}{\Gamma_{n,m}}\left[\sum_{m'\in C_{n'}} e^{(\tilde{w}_{n',m'}+\tilde{\rho}_{n',m'}-\tilde{\rho}_{n,m})}\right]\right),$$

which is a log-sum-exp function and hence convex [19]. Therefore, the transformed problem (4.7) is convex.

Since Problem (4.7) is convex, we can use Lagrangian techniques to solve it. The Lagrangian function associated with (4.7) is:

$$\begin{aligned} L(\tilde{\boldsymbol{\rho}},\tilde{\boldsymbol{w}},\tilde{\boldsymbol{t}},\boldsymbol{\lambda},\boldsymbol{\mu},\boldsymbol{\nu}) = {} & \sum_{n\in\mathcal{N}}\sum_{m\in C_{n}} e^{\tilde{w}_{n,m}+\tilde{\rho}_{n,m}} + \sum_{n\in\mathcal{N}}\lambda_{n}\left[\sum_{m\in C_{n}} e^{\tilde{w}_{n,m}} - 1\right] \\ & + \sum_{n\in\mathcal{N}}\sum_{m\in C_{n}}\mu_{n,m}\left[\log c_{n,m}^{tar} - \tilde{w}_{n,m} - \log\log\left(1+e^{\tilde{t}_{n,m}}\right)\right] \\ & + \sum_{n\in\mathcal{N}}\sum_{m\in C_{n}}\nu_{n,m}\left[\tilde{t}_{n,m} - \log\left(\Lambda_{n,m}(\tilde{\boldsymbol{\rho}},\tilde{\boldsymbol{w}})\right)\right], \end{aligned}$$

where $\boldsymbol{\lambda}$, $\boldsymbol{\mu}$ and $\boldsymbol{\nu}$ are the Lagrangian multipliers with:

$$\begin{aligned}\boldsymbol{\lambda} &= (\lambda_n)_{n\in\mathcal{N}},\\ \boldsymbol{\mu} &= (\boldsymbol{\mu}_n)_{n\in\mathcal{N}} \quad \text{where} \quad \boldsymbol{\mu}_n = (\mu_{n,m})_{m\in C_n} \text{ and}\\ \boldsymbol{\nu} &= (\boldsymbol{\nu}_n)_{n\in\mathcal{N}} \quad \text{where} \quad \boldsymbol{\nu}_n = (\nu_{n,m})_{m\in C_n}.\end{aligned}$$

The dual problem associated with (4.7) is given by:

$$\max_{\boldsymbol{\lambda},\boldsymbol{\mu},\boldsymbol{\nu}} \; U(\boldsymbol{\lambda},\boldsymbol{\mu},\boldsymbol{\nu}) \tag{4.11}$$

where

$$U(\boldsymbol{\lambda},\boldsymbol{\mu},\boldsymbol{\nu}) = \min_{\tilde{\boldsymbol{\rho}},\tilde{\boldsymbol{w}},\tilde{\boldsymbol{t}}} \; L(\tilde{\boldsymbol{\rho}},\tilde{\boldsymbol{w}},\tilde{\boldsymbol{t}},\boldsymbol{\lambda},\boldsymbol{\mu},\boldsymbol{\nu}). \tag{4.12}$$

The partial derivatives of $\Lambda_{n',m'}(\boldsymbol{\rho},\boldsymbol{w})$, which we will make use of to derive the expressions for the partial derivatives of the Langrangian function, are given by:

$$\frac{\partial\Lambda_{n',m'}(\boldsymbol{\rho},\boldsymbol{w})}{\partial\rho_{n,m}} = \begin{cases}\Lambda_{n',m'}(\boldsymbol{\rho},\boldsymbol{w}), & \text{if } n'=n \text{ and } m'=m\\ 0, & \text{if } n'=n \text{ and } m'\neq m\\ -\Lambda^2_{n,m}(\boldsymbol{\rho},\boldsymbol{w})\frac{\Gamma_{n,m'}\rho_{n,m}w_{n,m}}{\Gamma_{n',m'}\rho_{n',m'}}, & \text{otherwise}\end{cases}$$

and

$$\frac{\partial\Lambda_{n',m'}(\boldsymbol{\rho},\boldsymbol{w})}{\partial w_{n,m}} = \begin{cases}0, & \text{if } n'=n\\ -\Lambda^2_{n,m}(\boldsymbol{\rho},\boldsymbol{w})\frac{\Gamma_{n,m'}\rho_{n,m}w_{n,m}}{\Gamma_{n',m'}\rho_{n',m'}}, & \text{otherwise.}\end{cases}$$

Writing out the partial derivatives of $L(\tilde{\boldsymbol{\rho}},\tilde{\boldsymbol{w}},\tilde{\boldsymbol{t}},\boldsymbol{\lambda},\boldsymbol{\mu},\boldsymbol{\nu})$, and then making the substitution $\rho_{n,m} = e^{\tilde{\rho}_{n,m}}$, $w_{n,m} = e^{\tilde{w}_{n,m}}$ and $t_{n,m} = e^{\tilde{t}_{n,m}}$, we obtain:

$$\frac{\partial L}{\partial\tilde{\rho}_{n,m}} = w_{n,m}\rho_{n,m}(1+\Phi_n(\tilde{\boldsymbol{\rho}},\tilde{\boldsymbol{w}},\boldsymbol{\nu})) - \nu_{n,m} \tag{4.13}$$

$$\frac{\partial L}{\partial\tilde{w}_{n,m}} = w_{n,m}\rho_{n,m}(1+\Phi_n(\tilde{\boldsymbol{\rho}},\tilde{\boldsymbol{w}},\boldsymbol{\nu})) - \mu_{n,m} + \lambda_n w_{n,m} \tag{4.14}$$

$$\frac{\partial L}{\partial \tilde{t}_{n,m}} = \frac{\mu_{n,m} t_{n,m}}{(1+t_{n,m})\log(1+t_{n,m})} + \nu_{n,m}, \tag{4.15}$$

where

$$\Phi_n(\tilde{\boldsymbol{\rho}}, \tilde{\boldsymbol{w}}, \boldsymbol{\nu}) = \sum_{n' \neq n} \sum_{m' \in \mathcal{C}_{n'}} \frac{\Gamma_{n,m'} \nu_{n',m'}}{\Gamma_{n',m'} \rho_{n',m'}} \Lambda_{n',m'}(\tilde{\boldsymbol{\rho}}, \tilde{\boldsymbol{w}}).$$

By setting the partial derivatives to zero, we get the expression for $\tilde{\boldsymbol{w}}$ that minimizes $L(\tilde{\boldsymbol{\rho}}, \tilde{\boldsymbol{w}}, \tilde{\boldsymbol{t}}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{\nu})$ in terms of the dual variables:

$$w_{n,m} = \frac{\mu_{n,m} - \nu_{n,m}}{\lambda_n}. \tag{4.16}$$

However, the corresponding expressions for $\tilde{\boldsymbol{\rho}}$ and $\tilde{\boldsymbol{t}}$ cannot be derived from the dual variables alone. Consequently, a closed-form expression for $U(\boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{\nu})$ (in Eq. (4.12)) cannot be found.

Therefore, Problem (4.11) can only be solved numerically with an iterative approach. To help with the development of this iterative algorithm, we first consider the following problem:

$$\arg \min_{\tilde{\boldsymbol{\rho}}, \tilde{\boldsymbol{w}}, \tilde{\boldsymbol{t}}} L(\tilde{\boldsymbol{\rho}}, \tilde{\boldsymbol{w}}, \tilde{\boldsymbol{t}}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{\nu}), \tag{4.17}$$

where $\boldsymbol{\lambda}, \boldsymbol{\mu}$ and $\boldsymbol{\nu}$ are given. We note that for given $\boldsymbol{\lambda}, \boldsymbol{\mu}$ and $\boldsymbol{\nu}$, the Lagrangian function $L(\tilde{\boldsymbol{\rho}}, \tilde{\boldsymbol{w}}, \tilde{\boldsymbol{t}}, \boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{\nu})$ is convex in the primal variables $\tilde{\boldsymbol{\rho}}, \tilde{\boldsymbol{w}}$ and $\tilde{\boldsymbol{t}}$. Therefore, Problem (4.17) can be solved using the gradient descent method [13] as described below.

Algorithm 4.4.1.

- ***Initialization:*** *Start with initial values for the primal variables:* $\tilde{\boldsymbol{\rho}}^{(0)}$, $\tilde{\mathbf{w}}^{\dagger}$ *and* $\tilde{\mathbf{t}}^{(0)}$. *Note that the value of* $\tilde{\mathbf{w}}^{\dagger}$ *can be computed using* (4.16). *Set* $l = 0$.

- ***Iteration*** l***:***

  Compute $\tilde{\boldsymbol{\rho}}^{(l+1)}$ *and* $\tilde{\mathbf{t}}^{(l+1)}$ *with gradient descent method (the descent directions are*

given by the respective gradients (4.13) *and* (4.15)*):*

$$\tilde{\rho}_{n,m}^{(l+1)} = \tilde{\rho}_{n,m}^{(l)} - \epsilon \left[w_{n,m}^{\dagger} \rho_{n,m}^{(l)} \left\{ 1 + \Phi_n \left(\tilde{\boldsymbol{\rho}}^{(l)}, \tilde{\mathbf{w}}^{\dagger}, \boldsymbol{\nu} \right) \right\} - \nu_{n,m} \right]$$

$$\tilde{t}_{n,m}^{(l+1)} = \tilde{t}_{n,m}^{(l)} - \epsilon \left[\frac{\mu_{n,m}\, t_{n,m}^{(l)}}{\left(1 + t_{n,m}^{(l)}\right) \log\left(1 + t_{n,m}^{(l)}\right)} + \nu_{n,m} \right]$$

where ϵ is a sufficiently small step size.

Note that each iteration of Algorithm 4.4.1 involves computing $\Phi_n(\ldots)$ which requires information from other cell users. This information can only be acquired by explicit cooperation between cells, for example, via message passing [24].

With the aid of Algorithm 4.4.1, Problem (4.11) can now be solved using an iterative algorithm. The basic idea of this prospective algorithm is as follows. At each iteration of the prospective algorithm, the optimal values of the primal variables are determined by solving (4.17) for the given values of the dual variables. These values are then used to update the values of the dual variables with the gradient descent method [13]. The algorithm is described below.

Algorithm 4.4.2.

- *Initialize the dual variables: $\boldsymbol{\lambda}^{(0)}, \boldsymbol{\mu}^{(0)}$ and $\boldsymbol{\nu}^{(0)}$. Set $k = 0$.*
- *Iteration k:*
  - *Given $\boldsymbol{\lambda}^{(k)}, \boldsymbol{\mu}^{(k)}$ and $\boldsymbol{\nu}^{(k)}$, solve* (4.17) *using Algorithm 4.4.1 in order to find the optimal values $(\tilde{\boldsymbol{\rho}}^{(k)}, \tilde{\mathbf{w}}^{(k)}, \tilde{\mathbf{t}}^{(k)})$.*
  - *Using the computed values $(\tilde{\boldsymbol{\rho}}^{(k)}, \tilde{\mathbf{w}}^{(k)}, \tilde{\mathbf{t}}^{(k)})$, update the dual variables using gradient descent:*

$$\lambda_n^{(k+1)} = \left[\lambda_n^{(k)} + \epsilon' \left\{ \sum_{m \in C_n} w_{n,m}^{(k)} - 1 \right\} \right]$$

$$\mu_{n,m}^{(k+1)} = \left[\mu_{n,m}^{(k)} + \epsilon' \left\{ \log c_{n,m}^{tar} - \tilde{w}_{n,m}^{(k)} - \log\left(1 + t_{n,m}^{(k)}\right) \right\} \right]^{+}$$

$$\nu_{n,m}^{(k+1)} = \left[\nu_{n,m}^{(k)} + \epsilon' \left\{ \tilde{t}_{n,m}^{(k)} - \log\left(\Lambda_{n,m}\left(\tilde{\boldsymbol{\rho}}^{(k)}, \tilde{\mathbf{w}}^{(k)} \right) \right) \right\} \right]^{+}$$

where $x^+ = \max(0, x)$ *and* ϵ' *is a sufficiently small step size.*

Each iteration of Algorithm 4.4.2 solves an inner minimization (4.17) for given values of the dual variables. This inner minimization is solved by Algorithm 4.4.1, each iteration of which involves computing the value of $\Phi_n(\ldots)$. Since the value of $\Phi_n(\ldots)$ depends on the parameters from other cells, evaluating it requires explicit cooperation between different cells. Therefore, this approach will not lead to a distributed solution.

4.5 Distributed Solution: Scheme 1

Algorithm 4.4.2 developed in Section 4.4 to solve Problem (4.5) requires explicit cooperation between cells. In this section, we devise a distributed algorithm which does not require explicit cooperation between cells to solve (4.5). This distributed algorithm will be executed by each cell to determine the transmit power and subchannel allocation for its users based only on the local measurements of its users.

The approach that we take in this section to solve Problem (4.5) is motivated by the fact that, in order to minimize the aggregate transmit powers in the network, the transmit power in each cell must be minimized. Thus, we can adopt a strategy in which the original problem (4.5) is decomposed into a set of N single cell subproblems, each of which is convex. These subproblems form the basis of an iterative algorithm. At each iteration of the algorithm, each cell solves a local power minimization problem with the assumption that the other cells are not changing their powers. The convergence of this algorithm is guaranteed by Yates framework [106].

4.5.1 Subproblem for Cell n

The subproblem for cell n corresponds to a problem of finding the optimal transmit PSD and weight vectors $(\boldsymbol{\rho}_n^*(\boldsymbol{q}), \boldsymbol{w}_n^*(\boldsymbol{q}))$ for the cell, that minimizes the total transmit power in the cell while meeting the rate requirements of its users when

the vector of total transmit powers for the network, excluding n is given by $\boldsymbol{q}\backslash q_n$, the components of $\boldsymbol{q}$ excluding q_n.

For each cell $n \in \mathcal{N}$, define the convex set $\boldsymbol{\Omega}_n$ by

$$\boldsymbol{\Omega}_n = \left\{ \boldsymbol{w}_n : \sum_{m \in C_n} w_{n,m} = 1 \text{ and } w_{n,m} > 0, \forall m \in C_n \right\}, \tag{4.18}$$

which are the available weight vectors for base station n. The optimal vectors of power densities and weights $(\boldsymbol{\rho}_n^*(\boldsymbol{q}), \boldsymbol{w}_n^*(\boldsymbol{q}))$ to use at base station n under $\boldsymbol{q}$ will solve the optimization problem:

$$\min_{\boldsymbol{\rho}_n, \boldsymbol{w}_n \in \boldsymbol{\Omega}_n} \quad \sum_{m \in C_n} w_{n,m}\, \rho_{n,m} \tag{4.19}$$

subject to

$$\begin{aligned} w_{n,m}\, \log\left(1 + \gamma_{n,m}(\rho_{n,m}, \boldsymbol{q})\right) &\geq c_{n,m}^{tar},\ \forall m \in C_n, \\ \rho_{n,m} &> 0,\ \forall\, m \in C_n. \end{aligned}$$

Rewriting the optimization (4.19) in terms of $p_{n,m}$'s, we obtain:

$$\min_{\boldsymbol{p}_n, \boldsymbol{w}_n \in \boldsymbol{\Omega}_n} \quad \sum_{m \in C_n} p_{n,m} \tag{4.20}$$

subject to

$$\begin{aligned} w_{n,m} \log\left(1 + \frac{1}{K_{n,m}(\boldsymbol{q})} \frac{p_{n,m}}{w_{n,m}}\right) &\geq c_{n,m}^{tar},\ \forall m \in C_n, \\ p_{n,m} &> 0, \quad \forall\, m \in C_n, \end{aligned} \tag{4.21}$$

where

$$K_{n,m}(\boldsymbol{q}) \;=\; \frac{1}{\Gamma_{n,m}} \left[\sigma_m^2 + \sum_{k \neq n} \Gamma_{k,m}\, q_k \right] \tag{4.22}$$

is fixed since $\boldsymbol{q}\backslash q_n$ is given. Note that $K_{n,m}(\boldsymbol{q})$ is monotonically increasing in $\boldsymbol{q}$.

By rearranging the rate constraints (4.21), we obtain:

$$p_{n,m} \;\geq\; \beta_{n,m}(w_{n,m}) \;\equiv\; K_{n,m}(\boldsymbol{q})\, w_{n,m} \left[\exp\left(\frac{c_{n,m}^{tar}}{w_{n,m}}\right) - 1\right], \quad \forall m \in C_n. \tag{4.23}$$

Note that since $K_{n,m}(\boldsymbol{q})$ is fixed, we do not explicitly show the dependence of $\beta_{n,m}$ on $\boldsymbol{q}$. From (4.23), it is clear that $\beta_{n,m}(w_{n,m})$ is the minimum transmit power required to satisfy the normalized rate requirement of user m, when the proportion of bandwidth allocated to that user is $w_{n,m}$.

Lemma 4.5.1. *$\beta_{n,m}(w_{n,m})$ is strictly convex for $w_{n,m} > 0$.*

Proof. We use the second order condition [19] for this purpose. Writing out the first and second derivatives of $\beta_{n,m}(w_{n,m})$, we have:

$$\begin{aligned}
\beta'_{n,m}(w_{n,m}) &= -K_{n,m}(\boldsymbol{q}) \left\{1 + \left[\frac{c_{n,m}^{tar}}{w_{n,m}} - 1\right] \exp\left(\frac{c_{n,m}^{tar}}{w_{n,m}}\right)\right\}, \\
\beta''_{n,m}(w_{n,m}) &= K_{n,m}(\boldsymbol{q})\, \frac{\left(c_{n,m}^{tar}\right)^2}{w_{n,m}^3} \exp\left(\frac{c_{n,m}^{tar}}{w_{n,m}}\right).
\end{aligned}$$

Clearly, $\beta''_{n,m}(w_{n,m}) > 0$ since $w_{n,m} > 0$. Hence $\beta_{n,m}(w_{n,m})$ is strictly convex for $w_{n,m} > 0$.

□

Using (4.23), we can write Problem (4.20) as a weight optimization problem:

$$\min_{\boldsymbol{w}_n \in \boldsymbol{\Omega}_n} \sum_{m \in C_n} \beta_{n,m}(w_{n,m}). \tag{4.24}$$

Since $\beta_{n,m}(w_{n,m})$ is convex for $w_{n,m} > 0$ and Ω_n is a convex set, Problem (4.24) is convex and hence has a unique solution.

Since Problem (4.24) is convex, it can be solved using Lagrangian techniques locally at base station n, given a measurement of $K_{n,m}(\boldsymbol{q})$. Writing the Lagrangian

associated with (4.24), we obtain:

$$\begin{aligned} L_n(\boldsymbol{w}_n, \lambda_n) &= \sum_{m \in C_n} K_{n,m}(\boldsymbol{q})\, w_{n,m} \left[\exp\left(\frac{c_{n,m}^{tar}}{w_{n,m}} \right) - 1 \right] + \lambda_n \left[\sum_{m \in C_n} w_{n,m} - 1 \right] \\ &= \sum_{m \in C_n} w_{n,m} \left\{ \lambda_n + K_{n,m}(\boldsymbol{q}) \left[\exp\left(\frac{c_{n,m}^{tar}}{w_{n,m}} \right) - 1 \right] \right\} - \lambda_n. \end{aligned}$$

where λ_n is the associated Lagrangian multiplier.

We use Karush-Khun-Tucker (KKT) conditions [19] to solve Problem (4.24). The primal, dual optimal points $(\boldsymbol{w}_n^*, \lambda_n^*)$ of (4.24) satisfy:

$$\lambda_n = K_{n,m}(\boldsymbol{q}) \left\{ 1 + \left[\frac{c_{n,m}^{tar}}{w_{n,m}} - 1 \right] \exp\left(\frac{c_{n,m}^{tar}}{w_{n,m}} \right) \right\}, \quad \forall m \in C_n, \tag{4.25}$$

$$\sum_{m \in C_n} w_{n,m} = 1. \tag{4.26}$$

Now consider the following function:

$$b_{n,m}(w_{n,m}) = K_{n,m}(\boldsymbol{q}) \left\{ 1 + \left[\frac{c_{n,m}^{tar}}{w_{n,m}} - 1 \right] \exp\left(\frac{c_{n,m}^{tar}}{w_{n,m}} \right) \right\}. \tag{4.27}$$

Using (4.27), Eq. (4.25) - (4.26) can be rewritten as:

$$b_{n,m}(w_{n,m}) = \lambda_n, \quad \forall m \in C_n, \tag{4.28}$$

$$\sum_{m \in C_n} w_{n,m} = 1. \tag{4.29}$$

Note that $b_{n,m}(w_{n,m})$ is monotonically decreasing in $w_{n,m} \in \mathbb{R}_+$ since $\frac{db_{n,m}}{dw_{n,m}} < 0$ for $w_{n,m} > 0$. We also have that $b_{n,m}(0^+) = \infty$ and $b_{n,m}(\infty) = 0$ (see Figure 4.2 for an illustration). For any $\lambda_n \geq 1$, the equation $b_{n,m}(w_{n,m}) = \lambda_n$ has a unique solution that is positive. On the other hand, for any $0 < \lambda_n < 1$, the equation $b_{n,m}(w_{n,m}) = \lambda_n$ has two solutions, one positive and the other negative. Consequently, for any $\lambda_n > 0$, the set of simultaneous equations $b_{n,m}(w_{n,m}) = \lambda_n$ for all $m \in C_n$, has exactly one solution $\bar{\boldsymbol{w}}_n(\lambda_n)$ with the property that $\bar{w}_{n,m}(\lambda_n) > 0$

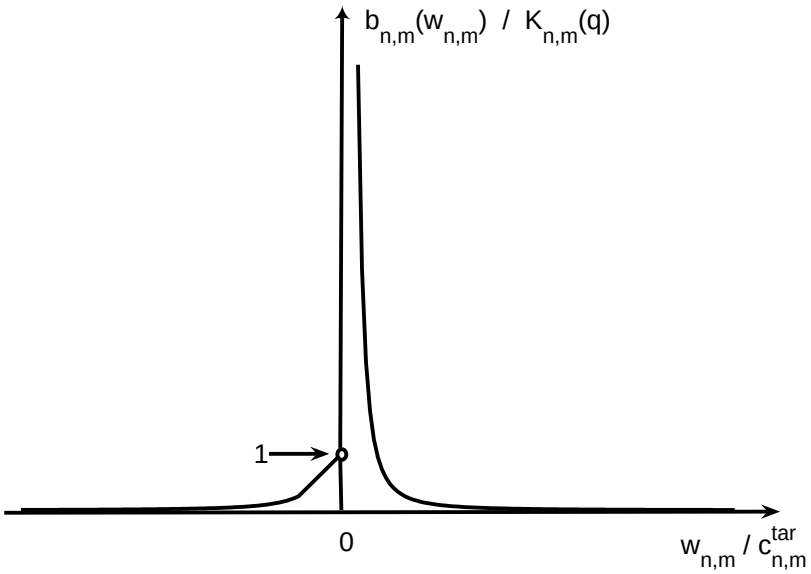


Figure 4.2: Plot of $b_{n,m}(w_{n,m})\ / K_{n,m}(\boldsymbol{q})$ vs. $w_{n,m}\ /\ c^{tar}_{n,m}$, where $K_{n,m}(\boldsymbol{q})$ and $c^{tar}_{n,m}$ are fixed. Note that there is a discontinuity at 0. We are only interested in the case where $w_{n,m} > 0$.

for all $m \in C_n$. Therefore, the problem of solving (4.28) - (4.29) reduces to a problem of searching for a $\lambda_n > 0$ (and the corresponding $\bar{\boldsymbol{w}}_n(\lambda_n) > 0$) that satisfies $\sum_{m \in C_n} \bar{w}_{n,m}(\lambda_n) = 1$.

Note that $\bar{w}_{n,m}(\lambda_n)$ is monotonically decreasing in λ_n since $\bar{w}_{n,m}(\lambda_n) > 0$ is a solution to $b_{n,m}(w_{n,m}) = \lambda_n$ and $b_{n,m}(w_{n,m})$ is monotonically decreasing for $w_{n,m} > 0$ (refer to Figure 4.2). Consequently, it follows that $\sum_{m \in C_n} \bar{w}_{n,m}(\lambda_n)$ is also monotonically decreasing in λ_n. Therefore, the solution λ_n^* of $\sum_{m \in C_n} \bar{w}_{n,m}(\lambda_n) = 1$ can be determined by a bisection search [10], if we can find a lower and an upper bound to λ_n^*, i.e., find any λ_n^l and λ_n^u such that $\lambda_n^l \leq \lambda_n^* \leq \lambda_n^u$. The following lemma provides a simple way of choosing these bounds.

Lemma 4.5.2. *If* $\mathbf{w}_n \in \boldsymbol{\Omega}_n$, *then,*

$$\min_{m \in C_n} b_{n,m}(w_{n,m}) \ \leq \ \lambda_n^* \ \leq \ \max_{m \in C_n} b_{n,m}(w_{n,m}). \tag{4.30}$$

Proof. Since $\sum_{m \in C_n} w_{n,m} = \sum_{m \in C_n} \bar{w}_{n,m}(\lambda_n^*) = 1$, there exists an m' such that $w_{n,m'} \leq \bar{w}_{n,m'}(\lambda_n^*)$ and another m'' such that $w_{n,m''} \geq \bar{w}_{n,m''}(\lambda_n^*)$. The required result follows from the fact that $b_{n,m}(\cdot)$ is monotonically decreasing in $\mathbb{R}_+$ and

$b_{n,m}(\bar{w}_{n,m}(\lambda_n^*)) = \lambda_n^*$ for all $m \in C_n$. □

Once λ_n^* is determined, the optimal $\boldsymbol{w}_n^*(\boldsymbol{q})$ and $\boldsymbol{p}_n^*(\boldsymbol{q})$ can consequently be computed since $w_{n,m}^*(\boldsymbol{q}) = \bar{w}_{n,m}(\lambda_n^*), \forall m \in C_n$ and $p_{n,m}^*(\boldsymbol{q}) = \beta_{n,m}(w_{n,m}^*(\boldsymbol{q})), \forall m \in C_n$.

We emphasize that the solution $(\boldsymbol{p}_n^*(\boldsymbol{q}), \boldsymbol{w}_n^*(\boldsymbol{q}))$ is the solution under the assumption that the interference in the cell of interest here (cell n) is generated by the other cells, when their transmit powers are fixed, as dictated by the fixed vector $\boldsymbol{q}$ of cell transmit powers. In the next section we address the fact that the cells are in reality coupled, and the transmit powers in the other cells are themselves determined by solving similar optimization problems to that described in the present section.

4.5.2 Iterative Algorithm

The minimum power required for user $m \in C_n$ to meet its normalized rate target is $p_{n,m}^*(\boldsymbol{q})$ when the interference is generated by a vector $\boldsymbol{q}$ of cell transmit powers. Therefore, the minimum cell transmit power required in cell n to satisfy all users $m \in C_n$ under $\boldsymbol{q}$ is $I_n(\boldsymbol{q}) = \sum_{m \in C_n} p_{n,m}^*(\boldsymbol{q})$. If the total power at base station n is at least $I_n(\boldsymbol{q})$, then it is possible to find feasible power and weight allocations that can support all users in the cell. Consequently, the necessary and sufficient condition for the existence of a $\boldsymbol{w}_n$ such that base station n can support all of its users is:

$$q_n \;\geq\; I_n(\boldsymbol{q}).$$

The requirement for the network to be able to support the rate requirements of all users is given by the vector inequality:

$$\boldsymbol{q} \;\geq\; \boldsymbol{I}(\boldsymbol{q})$$

where

$$\boldsymbol{I}(\boldsymbol{q}) = (I_n(\boldsymbol{q}))_{n \in \mathcal{N}}. \tag{4.31}$$

Note that, throughout this thesis, cell vector inequalities are defined componentwise. We refer to $\boldsymbol{I}(\boldsymbol{q})$ as the interference function to be consistent with Yates framework in [106]. To show that the solution to the global multi-cell problem (4.5) is unique, we apply Yates framework [106]. Central to Yates framework [106] is the notion of a *standard interference function*. A standard interference function is defined in [106] as follows:

Definition 4.5.1. An interference function $\boldsymbol{I}(\boldsymbol{q})$ is *standard* if for all $\boldsymbol{q} \geq 0$ the following properties are satisfied[1]:

- *Positivity:* $\boldsymbol{I}(\boldsymbol{q}) > 0$ for all $\boldsymbol{q} \geq 0$.
- *Monotonicity:* If $\boldsymbol{q} \geq \boldsymbol{q}'$, then $\boldsymbol{I}(\boldsymbol{q}) \geq \boldsymbol{I}(\boldsymbol{q}')$.
- *Scalability:* For all $\alpha > 1$, $\alpha \boldsymbol{I}(\boldsymbol{q}) > \boldsymbol{I}(\alpha \boldsymbol{q})$.

Note that the vector inequality $\boldsymbol{q} > \boldsymbol{q}'$ corresponds to a strict inequality in all components.

Lemma 1 in [106] states that if $\boldsymbol{I}(\boldsymbol{q})$ is *standard*, and if there is any feasible power vector $\boldsymbol{q}$ satisfying $\boldsymbol{q} \geq \boldsymbol{I}(\boldsymbol{q})$ then $\boldsymbol{I}(\boldsymbol{q})$ has a unique fixed point, which is the *minimal* solution to the inequality $\boldsymbol{q} \geq \boldsymbol{I}(\boldsymbol{q})$.

In order to be able to apply Yates framework to the problem at hand, we show that the function $\boldsymbol{I}(\boldsymbol{q})$ defined in (4.31) is *standard*.

Theorem 4.5.3. $\mathbf{I}(\mathbf{q})$ *in* (4.31) *is standard.*

Proof.

- *Positivity*: this follows since $\sigma_m^2 > 0$.
- *Monotonicity*: Suppose that $\boldsymbol{q} \geq \boldsymbol{q}'$. When the vector of total powers for the network is $\boldsymbol{q}$, the following holds for user $m \in C_n$:

$$p_{n,m}^*(\boldsymbol{q}) \;=\; K_{n,m}(\boldsymbol{q})\; w_{n,m}^*(\boldsymbol{q}) \left[\exp\left(\frac{c_{n,m}^{tar}}{w_{n,m}^*(\boldsymbol{q})} \right) - 1 \right]. \tag{4.32}$$

[1] Leung *et al.* [55] show that the Positivity condition is redundant and, it is just a consequence of the other two conditions.

Now, if instead the interference to users in cell n was generated by the vector $\boldsymbol{q}'$, but the bandwidth allocated for user m remains $\boldsymbol{w}_n^*(\boldsymbol{q})$, the power allocation at base station n will be $\boldsymbol{p}'_n$, where

$$p'_{n,m} = K_{n,m}(\boldsymbol{q}')\, w^*_{n,m}(\boldsymbol{q}) \left[\exp\left(\frac{c^{tar}_{n,m}}{w^*_{n,m}(\boldsymbol{q})}\right) - 1\right]. \tag{4.33}$$

Using (4.32) and (4.33) and the fact that $K_{n,m}(\boldsymbol{q}) \geq K_{n,m}(\boldsymbol{q}')$, we have that $p^*_{n,m}(\boldsymbol{q}) \geq p'_{n,m}$. Since $\boldsymbol{w}^*_n(\boldsymbol{q})$ is not necessarily optimal when the interference is generated by $\boldsymbol{q}'$, we also have that $\sum_{m\in C_n} p'_{n,m} \geq \sum_{m\in C_n} p^*_{n,m}(\boldsymbol{q}')$. Hence, $\sum_{m\in C_n} p^*_{n,m}(\boldsymbol{q}) \geq \sum_{m\in C_n} p^*_{n,m}(\boldsymbol{q}')$, that is, $I_n(\boldsymbol{q}) \geq I_n(\boldsymbol{q}')$.

- *Scalability*: Suppose $\alpha > 1$. Let $\boldsymbol{p}'_n$ be the power allocation at base station n when it uses $\boldsymbol{w}^*_n(\boldsymbol{q})$ and the vector of total powers for the network is $\alpha\boldsymbol{q}$. Then,

$$p'_{n,m} = K_{n,m}(\alpha\boldsymbol{q})\, w^*_{n,m}(\boldsymbol{q}) \left[\exp\left(\frac{c^{tar}_{n,m}}{w^*_{n,m}(\boldsymbol{q})}\right) - 1\right]. \tag{4.34}$$

Using (4.32) and (4.34) and the fact that $\alpha K_{n,m}(\boldsymbol{q}) > K_{n,m}(\alpha\boldsymbol{q})$, we have that $\alpha p^*_{n,m}(\boldsymbol{q}) > p'_{n,m}$. Since $\boldsymbol{w}^*_n(\boldsymbol{q})$ is not necessarily optimal when the interference is generated by $\alpha\boldsymbol{q}$, we also have that $\sum_{m\in C_n} p'_{n,m} \geq \sum_{m\in C_n} p^*_{n,m}(\alpha\boldsymbol{q})$. Hence, $\sum_{m\in C_n} \alpha p^*_{n,m}(\boldsymbol{q}) > \sum_{m\in C_n} p^*_{n,m}(\alpha\boldsymbol{q})$, that is, $\alpha I_n(\boldsymbol{q}) > I_n(\alpha\boldsymbol{q})$.

□

Since $\boldsymbol{I}(\boldsymbol{q})$ is standard, the standard power control algorithm [106]:

$$\boldsymbol{q}^{(k+1)} = \boldsymbol{I}(\boldsymbol{q}^{(k)}) \tag{4.35}$$

can be used to solve Problem (4.5). Solving for $I_n(\boldsymbol{q})$ at base station n requires only the knowledge of $K_{n,m}(\boldsymbol{q}) = \rho_{n,m} \,/\, \gamma_{n,m}(\rho_{n,m}, \boldsymbol{q})$. Therefore, this algorithm can be executed locally at each base station with the feedback on achieved SIR from only its users, and hence is distributed.

Description of the Algorithm

The algorithm is summarized here.

Algorithm 4.5.1.

- ***Initialization:*** *Start with any initial allocation* $(\mathbf{p}_n^{(0)}, \mathbf{w}_n^{(0)})$ *and let* $q_n^{(0)} = \sum_{m \in C_n} p_{n,m}^{(0)}$. *Set* $k = 0$.

- ***Iteration*** k***:***

  1. *Solve Problem* (4.24) *for the optimal* $\mathbf{w}_n^*(\mathbf{q}^{(k)})$ *for the given interference generated by* $\mathbf{q}^{(k)}$ *(the value of* $K_{n,m}(\mathbf{q}^{(k)})$ *can be computed using the user SIR measurements).*
  2. *Compute* $\mathbf{p}_n^*(\mathbf{q}^{(k)})$ *using* (4.23)*:* $p_{n,m}^*(\mathbf{q}^{(k)}) = \beta_{n,m}(w_{n,m}^*(\mathbf{q}^{(k)}))$, $\forall m \in C_n$.
  3. *Use the computed values for the next iteration. Set* $\mathbf{p}_n^{(k+1)} = \mathbf{p}_n^*(\mathbf{q}^{(k)})$ *and* $\mathbf{w}_n^{(k+1)} = \mathbf{w}_n^*(\mathbf{q}^{(k)})$.

4.6 Discrete Allocation

The resource allocation problem (4.5) was formulated with continuous weights which allowed us to solve it using convex optimization tools. In this section, we deal with the corresponding problem with discrete weights.

The set of discrete values that a weight $w_{n,m}$ of user $m \in C_n$ can take, is given by the following set:

$$\mathbf{\Pi} = \left[\frac{1}{N_c}, \frac{2}{N_c}, \ldots, 1\right], \tag{4.36}$$

where N_c is the number of subcarriers in the system, i.e., $w_{n,m} \in \mathbf{\Pi}$. Using this definition, we can write down the resource allocation problem corresponding to the discrete weights:

$$\min_{\boldsymbol{\rho}, \boldsymbol{w}} \sum_{n \in \mathcal{N}} \sum_{m \in C_n} w_{n,m}\, \rho_{n,m} \tag{4.37}$$

such that for all n,

$$\begin{aligned} w_{n,m} \log\left(1+\gamma_{n,m}(\rho_{n,m},\boldsymbol{q})\right) &\geq c_{n,m}^{tar},\ \forall m \in C_n, \\ w_{n,m} &\in \mathbf{\Pi}, \quad \forall\, m \in C_n, \\ \sum_{m\in C_n} w_{n,m} &= 1, \\ \rho_{n,m} &> 0,\ \forall\, m \in C_n. \end{aligned}$$

Problem (4.37) is a mixed integer, nonlinear optimization problem where the variables $\rho_{n,m}$'s are continuous and $w_{n,m}$'s are discrete. Note that, in this thesis, we treat $\rho_{n,m}$'s as continuous variables. In practice, they might also be discrete. Before solving (4.37) exactly, we consider an approximate method.

4.6.1 Continuous Relaxation

One possible way to "solve" a discrete problem is by solving it as a continuous model and then rounding the continuous optimum [90]. Applying this method to the problem at hand, one first solves the continuous version of Problem (4.37). The continuous version of this problem is Problem (4.5) that we solved earlier. Once the solution is determined to the continuous problem, the weights can be rounded to the nearest discrete value.

It should be noted that rounding of weights to the nearest discrete value may not suffice as the sum of the discretized weights for each cell may no longer be 1. If that is the case, some modification is needed to the rounding process. One such modification corresponds to looking at the users for which the rounding error is greatest and rounding the other way until the sum of the discretized weights becomes 1.

After the discretization is performed, the transmit PSDs computed for the users with the continuous relaxation approach will no longer be accurate. The users for which the continuous weights are rounded up, will achieve better normalized rates than their targets. Conversely, the users for which the continuous weights are rounded down, will suffer. The simple rounding procedure described

above does not differentiate between users during the rounding of the weights. For a user with a smaller weight, rounding down will have a greater effect, as the proportion of loss is greater. The following rounding procedure reduces this effect by ensuring that rounding down is only done to users who can "afford" it. It starts with the list of weights to be discretized. The discretization of weights is done one at a time, starting with the smallest weight. The smallest weight in the list is rounded up to the nearest discrete value and is taken off the list. The increase in the sum due to this rounding up is then "distributed" among the weights that are still in the list so that the weights sum to unity again. This procedure is repeated until the list is empty.

We can also consider a variation to continuous relaxation that does not suffer from the effects of rounding down of the weights. This is achieved by allowing a "weight margin" so that the weights are always rounded up: i.e., the sum weight requirement is now changed to $\sum_{m \in C_n} w_{n,m} = 1 - \epsilon$ where an appropriate ϵ is chosen, e.g., $\epsilon = L_n / N_c$, which corresponds to reserving L_n subchannels. After the continuous version of the problem is solved, the continuous weights are now rounded up to the nearest discrete value. Even after this procedure, the sum of discretized weights may sum up to a value less than one, i.e., there may still be a few subchannels in the reserve. These unused subchannels can then be given to users based on some other criteria, e.g., to users with the least number of subchannels in order to improve their diversity.

The problem with this continuous relaxation approach is that even if a feasible solution can be obtained by rounding, such a solution may not be optimal or even close to optimal. This approach does not necessarily lead to to a good discrete solution. The rounding procedure at best may be regarded as a heuristic [90].

This leads us to the next section where we solve Problem (4.37) exactly.

4.6.2 Exact Solution

Problem (4.37) can be solved using an approach similar to the one we used in solving the continuous version of the problem (4.5) in Section 4.5. The Yates frame-

work can still be applied. The difference is in the way the optimal power and weight vectors for each cell are computed at each step of the power control algorithm. With Problem (4.5), each cell computes the optimal power and weight vectors by solving a convex optimization problem. With Problem (4.37), computing the optimal power and weight vectors involves solving a combinatorial optimization problem. The rest of the analysis will be exactly the same.

Subproblem for Cell n

For a cell $n \in \mathcal{N}$, define a set $\mathbf{\Pi}_n$ by

$$\mathbf{\Pi}_n \;=\; \left\{ \boldsymbol{w}_n : \sum_{m \in C_n} w_{n,m} = 1 \text{ and } w_{n,m} \in \mathbf{\Pi}, \forall m \in C_n \right\}, \tag{4.38}$$

which are the available weight vectors for base station n. Using this, we can write down the subproblem for cell n:

$$\min_{\boldsymbol{\rho}_n, \boldsymbol{w}_n \in \mathbf{\Pi}_n} \; \sum_{m \in C_n} w_{n,m}\, \rho_{n,m} \tag{4.39}$$

subject to

$$\begin{aligned} w_{n,m}\, \log\left(1 + \gamma_{n,m}(\rho_{n,m}, \boldsymbol{q})\right) &\geq c_{n,m}^{tar},\ \forall m \in C_n, \\ \rho_{n,m} &> 0,\ \forall\, m \in C_n. \end{aligned}$$

By applying the same techniques as in Section 4.5.1, Problem (4.39) becomes the following weight optimization problem:

$$\min_{\boldsymbol{w}_n \in \mathbf{\Pi}_n} \; \sum_{m \in C_n} \beta_{n,m}(w_{n,m}) \tag{4.40}$$

where $\beta_{n,m}(w_{n,m})$ is defined by (4.23). Problem (4.40) is combinatorial in $\boldsymbol{w}_n$. As we will show later in this section, Problem (4.40) can still be solved in polynomial time (i.e., without an exhaustive search) using the structure of the problem.

Towards this end, we first convert the weight allocation problem (4.40) into a

subchannel allocation problem by defining $\theta_{n,m}(\cdot)$ on $\mathbb{R}_+$ as:

$$\theta_{n,m}(x) \equiv \beta_{n,m}\left(\frac{x}{N_c}\right), \quad \forall m \in C_n.$$

Since $\beta_{n,m}(\cdot)$ is convex on $\mathbb{R}_+$, so is $\theta_{n,m}(\cdot)$.

Define $\boldsymbol{\psi} \equiv \{1, 2, \ldots, N_c\}$. Then, given $\eta_{n,m} \in \boldsymbol{\psi}$, $\theta_{n,m}(\eta_{n,m})$ represents the amount of power required to satisfy user m when the user is allocated $\eta_{n,m}$ subchannels. Furthermore, it can be easily verified that

$$\theta_{n,m}(\eta_{n,m}) \leq \theta_{n,m}(1), \qquad \forall \eta_{n,m} \in \boldsymbol{\psi},$$

i.e., a user requires the most amount of transmit power when only a single subchannel is allocated.

Define a set $\boldsymbol{\psi}_n$ for cell $n \in \mathcal{N}$:

$$\boldsymbol{\psi}_n = \left\{ (\eta_{n,m})_{m \in C_n} : \sum_{m \in C_n} \eta_{n,m} = N_c \text{ and } \eta_{n,m} \in \boldsymbol{\psi}, \forall m \in C_n \right\},$$

which are the available subchannel allocation vectors for cell n. Using this definition, we are now ready to write Problem (4.40) as a subchannel allocation problem:

$$\min_{\boldsymbol{\eta}_n \in \boldsymbol{\psi}_n} \sum_{m \in C_n} \theta_{n,m}(\eta_{n,m}). \tag{4.41}$$

We can quantify the reduction[2] in power due to an allocation of an additional subchannel for a user m who already has $\eta_{n,m}$ subchannels:

$$\delta_{n,m}(\eta_{n,m}) = \theta_{n,m}(\eta_{n,m}) - \theta_{n,m}(\eta_{n,m}+1), \qquad \eta_{n,m} = 1, \ldots, N_c - 1. \tag{4.42}$$

Note that since $\delta_{n,m}(\eta_{n,m})$ is defined for $1 \leq \eta_{n,m} \leq N_c - 1$, this notation implicitly assumes that each user gets allocated at least one subchannel. Since $\theta_{n,m}(x)$ is

[2]There is an abuse of terminology here. The quantity $\delta_{n,m}(\eta_{n,m})$ may not necessarily refer to a reduction in power, since $[\theta_{n,m}(\eta_{n,m}) - \theta_{n,m}(\eta_{n,m}+1)]$ may *not* be positive for all values of $\eta_{n,m} = 1, 2, \ldots, N_c - 1$.

convex in $x \in \mathbb{R}_+$, $\theta'_{n,m}(x)$ is monotonically increasing in $x \in \mathbb{R}_+$. Consequently,

$$\theta_{n,m}(a+1) - \theta_{n,m}(a)$$

is monotonically increasing in a for $a > 0$ and hence $\delta_{n,m}(\eta_{n,m})$ is monotonically decreasing in $\eta_{n,m}$.

Using (4.42), $\theta_{n,m}(\eta_{n,m})$ can be written as:

$$\theta_{n,m}(\eta_{n,m}) \;=\; \theta_{n,m}(1) - \sum_{i=1}^{\eta_{n,m}-1} \delta_{n,m}(i), \qquad \eta_{n,m} \in \boldsymbol{\psi}. \tag{4.43}$$

Note that when $\eta_{n,m} = 1$, the sum in (4.43) has no terms and is equal to zero. Using (4.43), we can write the subchannel allocation problem (4.41) as:

$$\begin{aligned}\min_{\boldsymbol{\eta}_n \in \boldsymbol{\psi}_n} \sum_{m \in C_n} \theta_{n,m}(\eta_{n,m}) \;&=\; \min_{\boldsymbol{\eta}_n \in \boldsymbol{\psi}_n} \sum_{m \in C_n} \left\{ \theta_{n,m}(1) - \sum_{i=1}^{\eta_{n,m}-1} \delta_{n,m}(i) \right\} \\ &=\; \sum_{m \in C_n} \theta_{n,m}(1) + \max_{\boldsymbol{\eta}_n \in \boldsymbol{\psi}_n} \sum_{m \in C_n} \sum_{i=1}^{\eta_{n,m}-1} \delta_{n,m}(i).\end{aligned}$$

Therefore, any subchannel allocation that solves (4.41) will also solve the following maximization problem and vice versa:

$$\max_{\boldsymbol{\eta}_n \in \boldsymbol{\psi}_n} \sum_{m \in C_n} \sum_{i=1}^{\eta_{n,m}-1} \delta_{n,m}(i). \tag{4.44}$$

There are $(N_c - L_n)$ terms within the sum in (4.44) where $L_n = |C_n|$ is the number of users in cell n (recall that each user automatically gets allocated one subchannel).

Now consider a matrix Δ_n of dimension $(N_c - 1) \times L_n$ where the (i, m) entry of the matrix has the value $\delta_{n,m}(i)$ as demonstrated in Figure 4.3. Moving down a column m (representing user m) of Δ_n, the values of $\delta_{n,m}(i)$ decrease since $\delta_{n,m}(\cdot)$ is a monotonically decreasing function. Hence, the first κ $(< N_c)$ entries of a particular column m are also the κ largest entries of that column. Consequently, the terms contributing to the maximum of (4.44) are also the $(N_c - L_n)$ largest

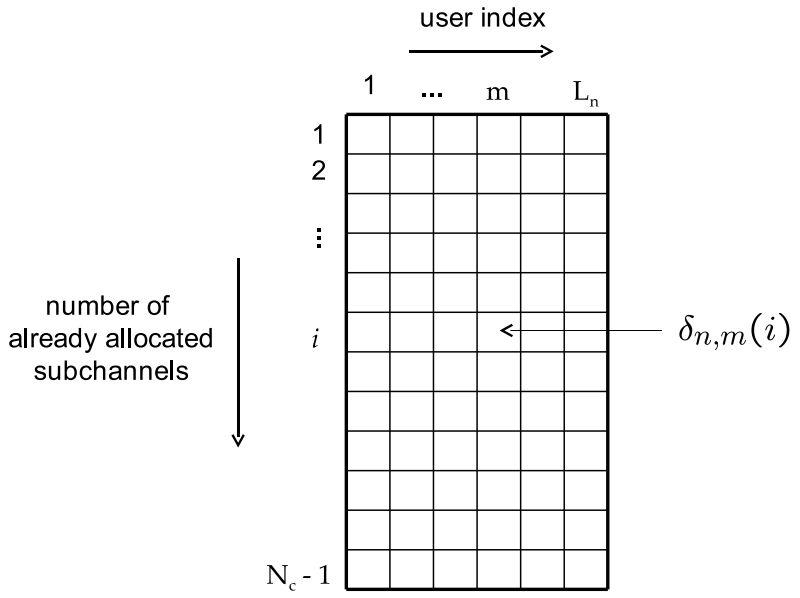


Figure 4.3: A diagrammatic illustration of matrix Δ_n. The rows correspond to the number of already allocated subchannels and the columns represent the users. The (i, m) entry has the value of $\delta_{n,m}(i)$, the reduction in power due to an allocation of an additional subchannel for user m who already has i subchannels. The value of $\delta_{n,m}(i)$ decreases as we go down a column.

entries of matrix Δ_n. By searching for these entries in Δ_n, an optimal subchannel allocation that solves (4.44) can be found.

We now describe an algorithm that exactly does that. To begin with, each user has exactly one subchannel allocated to it, i.e., $\boldsymbol{\eta}_n^{(0)} = \mathbf{1}$. The remaining $r_n = N_c - L_n$ subchannels are allocated to users, one at a time.

Let $\boldsymbol{\eta}_n^{(k)}$ be the subchannel allocation after $k + L_n$ subchannels have already been allocated (i.e., k subchannels have been allocated after the initial allocation of one subchannel per user). There are $r_n - k > 0$ remaining subchannels to be allocated. To determine the user to which the next subchannel is to be allocated, the algorithm examines the entries of Δ_n depending on the value of $\boldsymbol{\eta}_n^{(k)}$: for user m, the $(\eta_{n,m}^{(k)}, m)$ entry $\delta_{n,m}(\eta_{n,m}^{(k)})$ is considered. Allocating the next subchannel to user m' satisfies

$$m' \leftarrow \arg \max_{m \in \mathcal{C}_n} \delta_{n,m}(\eta_{n,m}^{(k)}).$$

The algorithm for solving (4.44) is formally stated here.

Algorithm 4.6.1.

- ***Initialization:*** *Start with one subchannel each per user, i.e.,* $\boldsymbol{\eta}_n^{(0)} = \mathbf{1}$.
- *Allocate* $r_n = N_c - L_n$ *subchannels to users, one at a time.*

for $k = 0, \ldots, r_n - 1$ ***do***
$\quad \boldsymbol{\eta}_n^{(k+1)} := \boldsymbol{\eta}_n^{(k)}$
$\quad m' \leftarrow \arg \max_{m \in C_n} \delta_{n,m}(\eta_{n,m}^{(k)})$
$\quad \eta_{n,m'}^{(k+1)} := \eta_{n,m'}^{(k)} + 1$
end for

Note that multiple subchannel allocations may solve Problem (4.44). Let $\boldsymbol{\eta}_n^*(\boldsymbol{q})$ be one such subchannel allocation and $\boldsymbol{w}_n^*(\boldsymbol{q})$ be the corresponding weight vector. The value of $\boldsymbol{p}_n^*(\boldsymbol{q})$ can then be determined by using (4.23).

By defining $I_n = \sum_{m \in C_n} p_{n,m}^*(\boldsymbol{q})$, it can be shown that the interference function $\boldsymbol{I}(\boldsymbol{q}) = (I_n(\boldsymbol{q}))_{n \in \mathcal{N}}$ is standard using exactly the same argument as used in Section 4.5. Consequently the standard power control algorithm can be used to find the optimal solution to Problem (4.37).

4.7 Conclusions

As part of solving a multipath frequency selective fading environment resource allocation problem in an OFDM cellular system using a two-layer approach, this chapter considered the higher layer resource allocation problem which is equivalent to a problem in a flat fading environment. The employment of the method of interference averaging allowed us to model the resource allocation to each user as a proportion of subchannels and power to that user, greatly simplifying the allocation problem. The problem was formulated as a power minimization, subject to meeting the user rate requirements. The treatment of the proportions of subchannels as continuous variables allowed us to apply various convex optimization tools such as Lagrangian techniques to the problem at hand.

We showed that the resource allocation problem is convex under logarithmic transformation. An iterative algorithm is derived to solve this transformed problem. However, this iterative algorithm requires explicit cooperation between cells and thus is not distributed.

Turning again to the original formulation of the problem, we developed an iterative algorithm which can be executed in a distributed manner by each base station with only the local information from the users within the cell. At each iteration of the algorithm, each cell solves a convex optimization problem to find the optimal resource allocation for its users using the local measurements taken by its users. This algorithm falls under Yates framework [106] that guarantees its convergence.

We also considered the discrete version of the higher layer resource allocation problem in which the proportions of subchannels are discrete. First, we considered approximate methods to solve the discrete problem. We then solved the discrete problem exactly. A provably convergent, distributed algorithm based on Yates framework was developed for this purpose.

We delay the presentation of numerical results until after the reduced complexity schemes have been considered in Chapter 5.

Chapter 5

Reduced Complexity Resource Allocation Under Flat Fading

5.1 Introduction

In this chapter, we continue our analysis of the higher layer resource allocation problem (4.5). In particular, we investigate the impact of reducing the complexity of the problem. This reduction in complexity is achieved by imposing additional constraints.

As we have noted earlier in Section 4.3, Problem (4.5) offers $2L_n - 1$ degrees of freedom to each base station n (L_n transmit PSDs and $L_n - 1$ weights) where L_n is the number of users in cell n. Imposing additional constraints on Problem (4.5) effectively reduces the number of degrees of freedom available to the base stations. We consider two such reduced complexity schemes in this chapter.

In Section 5.2, we study the "static bandwidth allocation" scheme which follows the "traditional" approach to resource allocation. The subchannel allocation to the users is static; the power control mechanism is then used to allocate the transmit powers to the users to meet their individual rate requirements. With this scheme, the number of degrees of freedom available to each base station n reduces to L_n from $2L_n - 1$. This problem is solved using the standard power control algorithm [106].

The second of the reduced complexity schemes considered is the "uniform transmit PSD allocation" scheme which constrains the transmit PSD to be the same for all users served by the base station (i.e., a flat transmit PSD across all subcarriers), but optimizes over the proportions of subcarriers allocated to each user. This scheme also yields L_n degrees of freedom to each base station n.

The motivation for studying the uniform transmit PSD allocation scheme is as follows. When we study the resource allocation in a frequency selective fading environment in Chapter 6, using the uniform PSD allocation scheme at the higher layer (together with fast frequency hopping) simplifies the design of the lower layer resource allocation algorithm. The lower layer resource allocation can now be performed independently in each cell, without knock-on effects between cells.

The uniform transmit PSD allocation problem is introduced in Section 5.3. Using Yates framework [106], the properties of its solution are characterized and the problem is solved with the standard power control algorithm. The drawback of using the standard power control algorithm in this case is that each iteration of the power control algorithm involves solving a nonlinear optimization problem. In Section 5.4, we devise a novel algorithm that avoids the shortcomings of the previous algorithm.

In section 5.5, we conduct extensive simulation studies to evaluate the three schemes that solve Problem (4.5). The discrete version of the uniform PSD allocation scheme is considered in Section 5.7 and a distributed algorithm is developed to solve it.

In Problem (4.5), $f(\gamma)$ is the spectral efficiency of a link with an SIR of γ. When we solved (4.5) exactly in Chapter 4, we used a particular f, namely, $f(\gamma) = \log(1+\gamma)$ based on the Shannon formula to make the analysis tractable. In contrast, we do not assume any specific choice for the function f when we solve the reduced complexity schemes in this chapter. We only require $f(\gamma)$ to be a continuous and monotonically increasing function of γ with $f(0) = 0$.

5.2 Static Bandwidth Allocation: Scheme 2

In this section, we consider the resource allocation problem (4.5) with static bandwidth allocation. This scheme incorporates the traditional approach to bandwidth allocation, such as used in [108, 34, 35, 106, 69], in the sense that the channel allocation to the users is fixed in advance; hence, so is the transmission quality requirement (e.g., signal to interference and noise ratio (SIR) threshold) of each link.

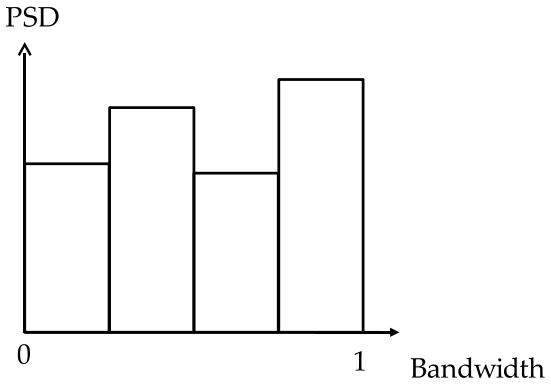


Figure 5.1: Static Bandwidth Allocation: The bandwidth allocation is proportional to the rate requirements of the users. The diagram here represents a case where the users have the same rate requirements. The resource allocation problem is to determine the transmit PSD values for users in order to satisfy their rate requirements.

The power control is then used to maintain the required transmission quality on each link.

With this scheme, the bandwidth (subchannel) allocation is statically performed based on the user rate requirements, and the transmit PSDs are then chosen for the users accordingly, in order to meet their rate requirements. The weight vectors are chosen a priori, proportional to the normalized rate requirements of the users. In cell n, the choice of the weight vector is given by:

$$w_{n,m} = \frac{c_{n,m}^{tar}}{\sum_{m' \in C_n} c_{n,m'}^{tar}}, \qquad \forall m \in C_n. \tag{5.1}$$

The resource allocation problem then is to determine the appropriate transmit power densities for users with the objective of power minimization, subject to meeting the rate requirements of the users. Refer to Figure 5.1 for an example. Mathematically, the static bandwidth allocation problem is:

$$\min_{\rho} \sum_{n \in \mathcal{N}} \sum_{m \in C_n} w_{n,m}\, \rho_{n,m} \tag{5.2}$$

such that for all n,

$$\begin{aligned} w_{n,m}\, f\left(\gamma_{n,m}(\rho_{n,m}, \boldsymbol{q})\right) &\geq c_{n,m}^{tar}, \quad \forall m \in C_n, \\ \rho_{n,m} &> 0, \qquad \forall\, m \in C_n, \end{aligned}$$

where the weights $w_{n,m}$ are chosen according to (5.1).

The number of degrees of freedom available to base station n is L_n: one transmit power density for each user in the cell. Because of the fact that the weights are chosen proportional to the normalized rate requirements, the SIR targets of the users within each cell will be equal by design:

$$\gamma_{n,m}^{tar} = f^{-1}\left(\sum_{m' \in C_n} c_{n,m'}^{tar}\right), \qquad \forall m \in C_n.$$

Consequently, the modulation scheme is the same across users in each cell.

Rewriting the rate constraint for user $m \in C_n$, we obtain the minimum transmit power density required to support that user, when the vector of total transmit powers for the network, excluding n, is given by $\boldsymbol{q}\backslash q_n$, the components of $\boldsymbol{q}$ excluding q_n:

$$\rho_{n,m} \geq \rho_{n,m}^{*}(\boldsymbol{q}) \equiv K_{n,m}(\boldsymbol{q}) \times \gamma_{n,m}^{tar}.$$

i.e., $\rho_{n,m}^{*}(\boldsymbol{q})$ is the minimum power density required to meet the normalized rate requirement of user m, when the interference is generated by $\boldsymbol{q}$. Accordingly, the necessary and sufficient condition for the existence of a power allocation for cell n to meet the normalized rate requirements of its users is:

$$q_n \geq I_n(\boldsymbol{q}) \equiv \sum_{m \in C_n} w_{n,m} \rho_{n,m}^{*}(\boldsymbol{q}).$$

It can be easily shown that $\boldsymbol{I}(\boldsymbol{q}) \equiv \left(I_n(\boldsymbol{q})\right)_{n \in \mathcal{N}}$ is standard [106], since $K_{n,m}(\boldsymbol{q})$ is standard and $I_n(\boldsymbol{q})$ is a linear combination of $K_{n,m}(\boldsymbol{q})$'s. Therefore, the standard power control algorithm can be used to solve Problem (5.2). At each step of the iteration, the power densities are selected to exactly meet the normalized rate

requirements of the users:

$$\boldsymbol{\rho}^{(k+1)} = \boldsymbol{\rho}^*(\boldsymbol{q}^{(k)}).$$

With the static subchannel allocation scheme, the number of subchannels for each user is allocated a priori, and thus done without taking account of the interference experienced by the user. The transmit power densities are then adapted based on the interference levels, to meet the normalized rate requirements of the users. As the users experiencing bad channel conditions require more power than that of the users with good channel conditions, this leads to a large dynamic range in the transmit power densities and consequently in the interference levels, requiring explicit interference averaging.

5.2.1 Static Bandwidth Allocation with Discrete Weights

The choice of the weights in the static bandwidth allocation scheme is made a priori and is independent of the power control algorithm. The discrete weight allocation can be obtained by rounding the continuous weights of (5.1) by using an approach similar to that of Section 4.6.1. The power control algorithm will consequently adapt the transmit power densities to suit these discretized weights.

This is in contrast to Scheme 1 (and the scheme to follow in Section 5.3) where the weight selection and the transmit power allocation are performed jointly.

5.3 Uniform PSD Allocation: Scheme 3

This section considers the resource allocation problem (4.5) with the additional constraint that each base station uses a flat transmit power spectrum, that is,

$$q_n = \rho_{n,m}, \qquad \forall m \in C_n, \tag{5.3}$$

since the bandwidth is taken to be 1.

Each base station can still choose an appropriate power level and vary the bandwidth allocation to its users. When a user has a poor SIR, it can receive more

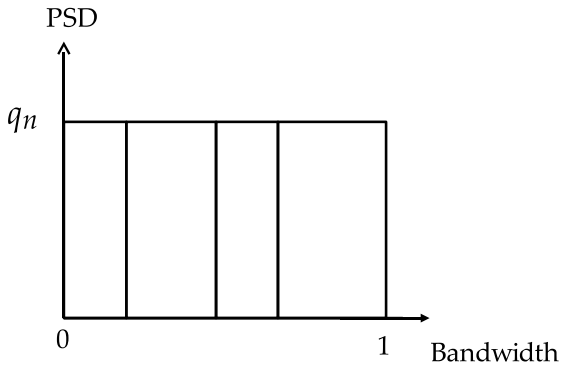


Figure 5.2: Uniform PSD Allocation: The transmit power for all users within a cell is constrained to be equal. The resource allocation problem is to determine a transmit PSD for the cell and a dynamic bandwidth allocation for each user in order to satisfy the normalized rate requirements of the users.

bandwidth. This implicitly allocates the user more received power. The number of degrees of freedom available to base station n is L_n, namely: $L_n - 1$ weights and a single transmit PSD. See Fig 5.2 for a demonstration.

Mathematically, the resource allocation problem corresponding to the uniform PSD allocation is:

$$\min_{\boldsymbol{q},\boldsymbol{w}} \sum_{n\in\mathcal{N}} q_n \tag{5.4}$$

such that for all n,

$$\left.\begin{aligned} w_{n,m}\, f\left(\gamma_{n,m}(q_n,\boldsymbol{q})\right) &\geq c_{n,m}^{tar},\ \forall m \in C_n, \\ \sum_{m\in C_n} w_{n,m} &= 1, \\ q_n &> 0, \\ w_{n,m} &> 0,\ \forall\, m \in C_n. \end{aligned}\right\} \tag{5.5}$$

Since the transmit PSD across all subcarriers is uniform in each cell, the interference averaging is realized automatically in a flat fading environment and explicit interference averaging based on fast frequency hopping is not necessary in this case.

Consider a fixed $\boldsymbol{w}_n$ for cell n that satisfies $\sum_{m\in C_n} w_{n,m} = 1$. We can now derive the expression for the minimum transmit power required by base station n

to meet the normalized rate requirements of all users in the cell, when the vector of transmit powers for the network, excluding n, is given by $\boldsymbol{q}\backslash q_n$ (i.e., when the interference is generated by $\boldsymbol{q}$). By rewriting the rate requirement for user $m \in C_n$, we obtain the constraint on the transmit power for base station n to be able to support user m:

$$q_n \;\geq\; J_{n,m}(\boldsymbol{q}, w_{n,m}) \equiv K_{n,m}(\boldsymbol{q}) \times f^{-1}\left(\frac{c_{n,m}^{tar}}{w_{n,m}}\right) \tag{5.6}$$

where $K_{n,m}(\boldsymbol{q})$ is given by (4.22). Thus, for base station n to be able to support all of its users $m \in C_n$, the constraint on its transmit power is:

$$q_n \;\geq\; \max_{m \in C_n} J_{n,m}(\boldsymbol{q}, w_{n,m}) \;\equiv\; J_n(\boldsymbol{q}, \boldsymbol{w}_n). \tag{5.7}$$

Correspondingly, the transmit power requirement for the network can be writing in the vector form as:

$$\boldsymbol{q} \;\geq\; \boldsymbol{J}(\boldsymbol{q}, \boldsymbol{w}), \tag{5.8}$$

where $\boldsymbol{J}(\boldsymbol{q}, \boldsymbol{w}) = (J_n(\boldsymbol{q}, \boldsymbol{w}_n))_{n \in \mathcal{N}}$.

5.3.1 Some Preliminary Results

We first develop some preliminaries that will assist with the characterization of the solution(s) of the uniform transmit PSD allocation problem (5.4), and which are used in the formulation and analysis of the distributed algorithms.

From (5.6), we have that, when the interference for user $m \in C_n$ is generated by the vector of transmit powers $\boldsymbol{q}$ (excluding q_n, since q_n is not relevant to the interference experienced by users in cell n) and the weight is $w_{n,m}$, then the minimum transmit power required to support user m is $J_{n,m}(\boldsymbol{q}, w_{n,m})$. Conversely, if we knew the minimum transmit power to support user m is t_n when the interference is generated by $\boldsymbol{q}$, then we can compute the required weight as a function

of t_n and $\boldsymbol{q}$ (from (5.6)):

$$\hat{w}_{n,m}(t_n, \boldsymbol{q}) \equiv \frac{c^{tar}_{n,m}}{f\left(\frac{t_n}{K_{n,m}(\boldsymbol{q})}\right)}. \tag{5.9}$$

Note that when $t_n = q_n$, (5.9) becomes,

$$\hat{w}_{n,m}(q_n, \boldsymbol{q}) = \frac{c^{tar}_{n,m}}{f\left(\gamma_{n,m}(q_n, \boldsymbol{q})\right)} \tag{5.10}$$

which is the proportion of subchannels required for user m when the power vector is $\boldsymbol{q}$.

Define $\sigma_n(t_n, \boldsymbol{q})$ as

$$\sigma_n(t_n, \boldsymbol{q}) \equiv \sum_{m \in C_n} \hat{w}_{n,m}(t_n, \boldsymbol{q}). \tag{5.11}$$

We refer to $\hat{\boldsymbol{w}}_n(t_n, \boldsymbol{q})$ as a pseudo weight vector to allude to the fact that it may not always be feasible (i.e., when the sum of weights $\sigma_n(t_n, \boldsymbol{q}) > 1$). Note that $\sigma_n(t_n, \boldsymbol{q})$ exhibits the following monotonicity properties:

- $\sigma_n(t_n, \boldsymbol{q})$ is monotonically decreasing in t_n.
- $\sigma_n(t_n, \boldsymbol{q})$ is monotonically increasing in $\boldsymbol{q}$.

Some additional properties of $\sigma_n(t_n, \boldsymbol{q})$ are contained in the following lemmas.

Lemma 5.3.1. *$\sigma_n(t_n, \mathbf{q}) \leq 1$ if and only if there is a feasible weight vector $\mathbf{w}_n$ such that the transmit power level of t_n is sufficient to satisfy each of the receivers $m \in C_n$ when the interference is generated by $\mathbf{q}$.*

Proof.

- *Sufficiency:* Suppose that $\sigma_n(t_n, \boldsymbol{q}) \leq 1$. Then, the weight vector $\hat{\boldsymbol{w}}_n(t_n, \boldsymbol{q})$ is feasible, and consequently, the transmit power level of t_n is sufficient to satisfy each of the users $m \in C_n$ when the weight vector is $\hat{\boldsymbol{w}}_n(t_n, \boldsymbol{q})$ and the interference is generated by $\boldsymbol{q}$.

- *Necessity:* Now suppose that there is a feasible weight vector $\boldsymbol{w}_n$ such that the transmit power level of t_n is sufficient to satisfy each of the users $m \in C_n$ when the interference is generated by $\boldsymbol{q}$. Then, $t_n \geq J_{n,m}(\boldsymbol{q}, w_{n,m})$, $\forall m \in C_n$. Consequently, $\hat{w}_{n,m}(t_n, \boldsymbol{q}) \leq w_{n,m}$, $\forall m \in C_n$. It immediately follows that $\sigma_n(t_n, \boldsymbol{q}) \leq 1$.

□

Lemma 5.3.2. *Given $\lambda > \mu$,*

$$\sigma_n(\lambda q_n, \lambda \mathbf{q}) \;\; < \;\; \sigma_n(\mu q_n, \mu \mathbf{q}). \tag{5.12}$$

Proof.
We have that $\gamma_{n,m}(\lambda q_n, \lambda \boldsymbol{q}) > \gamma_{n,m}(\mu q_n, \mu \boldsymbol{q})$ since $\lambda > \mu$. Using the fact that $f(\cdot)$ is a monotonically increasing function, we have that $\hat{w}_{n,m}(\lambda q_n, \lambda \boldsymbol{q}) < \hat{w}_{n,m}(\mu q_n, \mu \boldsymbol{q})$ (from (5.10)). The required result follows. □

Lemma 5.3.3. *$\pi(t_n) = \sigma_n(t_n, \mathbf{q}) - 1 = 0$ has a unique solution $\bar{t}_n$, and, for an arbitrary weight vector $\mathbf{w}_n \in \boldsymbol{\Omega}_n$, the following holds:*

$$\min_{m \in C_n} J_{n,m}(\mathbf{q}, w_{n,m}) \;\leq\; \bar{t}_n \;\leq\; \max_{m \in C_n} J_{n,m}(\mathbf{q}, w_{n,m}). \tag{5.13}$$

Proof.
Consider an arbitrary weight vector $\boldsymbol{w}_n \in \boldsymbol{\Omega}_n$. Let $t_{n,m} = J_{n,m}(\boldsymbol{q}, w_{n,m}), \forall m \in C_n$, $t_{\min} = \min_{m \in C_n} t_{n,m}$ and $t_{\max} = \max_{m \in C_n} t_{n,m}$. We make the following observations:

- A transmit power level of $t_{\min}$ is only sufficient to satisfy the user(s) that achieve the minimum in the definition of $t_{\min}$ above. Hence, $\sigma_n(t_{\min}, \boldsymbol{q}) \geq 1$ with equality if and only if $t_{\max} = t_{\min}$.

- A transmit power level of $t_{\max}$ is sufficient to satisfy each of the users $m \in C_n$ with $\boldsymbol{w}_n$. Therefore, using Lemma 5.3.1, $\sigma_n(t_{\max}, \boldsymbol{q}) \leq 1$ with equality if and only if $t_{\max} = t_{\min}$.

Since $\sigma_n(t_n, \boldsymbol{q})$ is a monotonically decreasing function of t_n, there exists a unique $\bar{t}_n$

such that $\sigma_n(\bar{t}_n, \boldsymbol{q}) = 1$ by the intermediate value theorem [36] and $t_{\min} \leq \bar{t}_n \leq t_{\max}$ (i.e., satisfying (5.13)). □

5.3.2 Characterization of the Minimal Solution

We use Yates monotonicity framework [106] to show that the minimization problem (5.4) has a unique solution. In order to do so, we first need to consider the problem of finding an optimal weight vector $\boldsymbol{w}$ for the network, given a fixed power allocation vector $\boldsymbol{q}$.

An optimal weight vector $\boldsymbol{w}_n^*(\boldsymbol{q})$ to use at base station n under $\boldsymbol{q}$ will solve the weight-optimization problem:

$$\min_{\boldsymbol{w}_n \in \boldsymbol{\Omega}_n} \; J_n(\boldsymbol{q}, \boldsymbol{w}_n), \tag{5.14}$$

where $\boldsymbol{\Omega}_n$ is a set of available weight vectors for base station n, given by (4.18). The uniqueness of $\boldsymbol{w}_n^*(\boldsymbol{q})$ is contained in the following lemma.

Lemma 5.3.4.

1. *Let $\bar{\mathbf{w}}_n$ be a weight vector that solves the optimization problem (5.14). Then, there exists a unique $\bar{t}_n$ such that, $J_{n,m}(\mathbf{q}, \bar{w}_{n,m}) = \bar{t}_n$, $\forall m \in C_n$.*

2. *The optimization (5.14) can be transformed into the equivalent problem of finding the unique $\bar{t}_n$ such that $\sigma_n(\bar{t}_n, \mathbf{q}) = 1$.*

3. *The weight vector $\bar{\mathbf{w}}_n$ that solves the optimization (5.14) is unique.*

Proof.

1. Let $\bar{t}_{n,m} = J_{n,m}(\boldsymbol{q}, \bar{w}_{n,m}), \forall m \in C_n$, $\bar{t}_{\max} = \max_{m \in C_n} \bar{t}_{n,m}$ and $\bar{t}_{\min} = \min_{m \in C_n} \bar{t}_{n,m}$. Suppose that $\bar{t}_{\max} \neq \bar{t}_{\min}$. Then, $\bar{t}_{\max} > \bar{t}_{\min}$. If the transmit power level at base station n is set to $\bar{t}_{\max}$, then, each user requires a fraction $\hat{w}_{n,m}(\bar{t}_{\max}, \boldsymbol{q})$ of subchannels to achieve the required rate requirement. Furthermore, we have that $\hat{w}_{n,m}(\bar{t}_{\max}, \boldsymbol{q}) \leq \hat{w}_{n,m}(\bar{t}_{n,m}, \boldsymbol{q}) = \bar{w}_{n,m}$ since $\bar{t}_{n,m} \leq \bar{t}_{\max}$ with a strict

inequality for at least one user. Therefore $\sigma_n(\bar{t}_{\max}, \boldsymbol{q}) < \sum_{m \in C_n} \bar{w}_{n,m} = 1$. Now define,

$$\bar{w}'_{n,m} = \frac{\hat{w}_{n,m}(\bar{t}_{\max}, \boldsymbol{q})}{\sigma_n(\bar{t}_{\max}, \boldsymbol{q})}, \quad \forall m \in C_n,$$

and note that $\bar{\boldsymbol{w}}'_n$ is feasible as it sums to unity. Since $\bar{w}'_{n,m} > \hat{w}_{n,m}(\bar{t}_{\max}, \boldsymbol{q})$ for all $m \in C_n$, we have that $J_{n,m}(\boldsymbol{q}, \bar{w}'_{n,m}) < \bar{t}_{n,m}$, for all $m \in C_n$. Consequently, $\max_{m \in C_n} J_{n,m}(\boldsymbol{q}, \bar{w}'_{n,m}) < \bar{t}_{\max}$, which contradicts the fact that $\bar{\boldsymbol{w}}_n$ solves (5.14). Therefore, $\bar{t}_{\max} = \bar{t}_{\min}$ and the result follows.

2. Let $\bar{\boldsymbol{w}}_n$ be the solution to (5.14) (as in 1) above). From 1), we have $\bar{w}_{n,m} = \hat{w}_{n,m}(\bar{t}_n, \boldsymbol{q}), \forall m \in C_n$. Therefore, $\sigma_n(\bar{t}_n, \boldsymbol{q}) = 1$ since $\sum_{m \in C_n} \bar{w}_{n,m} = 1$.

3. $\hat{w}_{n,m}(t_n, \boldsymbol{q})$ is unique for any given $t_n > 0$ since $\hat{w}_{n,m}(t_n, \cdot)$ is monotonically decreasing in t_n and $\boldsymbol{q}$ is fixed. Consequently, $\hat{\boldsymbol{w}}_n(t_n, \boldsymbol{q})$ is also unique for any given $t_n > 0$. Since $\bar{t}_n$ satisfying $\sigma_n(\bar{t}_n, \boldsymbol{q}) = 1$ is unique by Lemma 5.3.3, so is $\hat{\boldsymbol{w}}_n(\bar{t}_n, \boldsymbol{q})$. The required result follows since $\bar{\boldsymbol{w}}_n = \hat{\boldsymbol{w}}_n(\bar{t}_n, \boldsymbol{q})$.

□

One important implication of Lemma 5.3.4 (2) is that Problem (5.14) can be solved via solving a nonlinear equation $\sigma_n(\bar{t}_n, \boldsymbol{q}) = 1$ with a single variable $\bar{t}_n$. By Lemma 5.3.4 (3), we can define the optimal weight vector $\boldsymbol{w}_n^*(\boldsymbol{q})$ to use under $\boldsymbol{q}$:

$$\boldsymbol{w}_n^*(\boldsymbol{q}) \equiv \arg \min_{\boldsymbol{w}_n \in \boldsymbol{\Omega}_n} J_n(\boldsymbol{q}, \boldsymbol{w}_n).$$

Then, the requirement for the network to be able to support the required normalized rates is given by the vector inequality:

$$\boldsymbol{q} \geq \boldsymbol{J}(\boldsymbol{q}, \boldsymbol{w}^*(\boldsymbol{q})) \equiv \boldsymbol{I}(\boldsymbol{q}). \tag{5.15}$$

To show that the solution to Problem (5.4) is unique, we apply Yates framework [106]. In particular, we show that the function $\boldsymbol{I}(\boldsymbol{q})$ defined in (5.15) is standard [106].

Lemma 5.3.5. $\mathbf{I}(\mathbf{q})$ *defined in* (5.15) *is standard.*

Proof.

- *Positivity*: this follows since the noise $\sigma_m^2 > 0$.
- *Monotonicity*: If $\boldsymbol{q}' \leq \boldsymbol{q}$,

$$\begin{aligned} \boldsymbol{I}(\boldsymbol{q}') &= \boldsymbol{J}(\boldsymbol{q}', \boldsymbol{w}^*(\boldsymbol{q}')) \\ &\leq \boldsymbol{J}(\boldsymbol{q}', \boldsymbol{w}^*(\boldsymbol{q})) \\ &\leq \boldsymbol{J}(\boldsymbol{q}, \boldsymbol{w}^*(\boldsymbol{q})) = \boldsymbol{I}(\boldsymbol{q}) \end{aligned}$$

  where the first inequality stems from the fact that since $\boldsymbol{w}^*(\boldsymbol{q}')$ is optimal for $\boldsymbol{q}'$, any other $\boldsymbol{w}^*(\boldsymbol{q})$ will not decrease the value of $\boldsymbol{J}(\boldsymbol{q}', \cdot)$.

- *Scalability*: For all $\alpha > 1$,

$$\begin{aligned} \boldsymbol{I}(\alpha\boldsymbol{q}) &= \boldsymbol{J}(\alpha\boldsymbol{q}, \boldsymbol{w}^*(\alpha\boldsymbol{q})) \\ &\leq \boldsymbol{J}(\alpha\boldsymbol{q}, \boldsymbol{w}^*(\boldsymbol{q})) \\ &< \alpha\boldsymbol{J}(\boldsymbol{q}, \boldsymbol{w}^*(\boldsymbol{q})) = \alpha\boldsymbol{I}(\boldsymbol{q}) \end{aligned}$$

  where the second inequality stems from the fact that $K_{n,m}(\alpha\boldsymbol{q}) < \alpha K_{n,m}(\boldsymbol{q})$ for any $\alpha > 1$.

□

Since $\boldsymbol{I}(\boldsymbol{q})$ is standard, we can apply Theorem 1 and Lemma 1 in [106] to (5.15). Theorem 1 in [106] states that if $\boldsymbol{I}(\boldsymbol{q})$ is standard, and if it has a fixed point, then the fixed point is unique. Lemma 1 in [106] shows that if there is any feasible power vector $\boldsymbol{q}$ satisfying $\boldsymbol{q} \geq \boldsymbol{I}(\boldsymbol{q})$, then $\boldsymbol{I}(\boldsymbol{q})$ has a unique fixed point, which is the *minimal* solution to $\boldsymbol{q} \geq \boldsymbol{I}(\boldsymbol{q})$. The result of applying Theorem 1 and Lemma 1 of [106] to (5.15) is as follows.

Corollary 5.3.6. *If (5.5) is feasible, then there is a unique power allocation* $\mathbf{q}^*$ *which solves (5.4) and is minimal. Associated with* $\mathbf{q}^*$ *is the unique optimal weight vector* $\mathbf{w}^*(\mathbf{q}^*)$.

The following theorem provides a useful characterization of the minimal solution.

Theorem 5.3.7. *Suppose that (5.5) is feasible. Then, the following statements hold.*

1. *The function* $\mathbf{I}(\mathbf{q})$ *defined in (5.15) has a unique fixed point* $\mathbf{q}^*$*. Furthermore,*

$$\sigma_n(q_n^*, \mathbf{q}^*) \;=\; 1, \quad \forall n \in \mathcal{N}. \tag{5.16}$$

2. *The pair* $(\mathbf{q}^*, \mathbf{w}^*(\mathbf{q}^*))$ *is the unique solution to (5.4).*

3. *If a power vector* $\mathbf{q}^\dagger$ *satisfies,*

$$\sigma_n(q_n^\dagger, \mathbf{q}^\dagger) \;=\; 1, \quad \forall n \in \mathcal{N}, \tag{5.17}$$

then, $\mathbf{q}^\dagger = \mathbf{q}^*$.

Proof.

1. Consider any $\boldsymbol{q}$ that satisfies (5.5). Then, it directly follows from Lemma 1 of [106] that the function $\boldsymbol{I}(\boldsymbol{q}) = \boldsymbol{J}(\boldsymbol{q}, \boldsymbol{w}^*(\boldsymbol{q}))$ has a unique fixed point $\boldsymbol{q}^*$ $(= \boldsymbol{I}(\boldsymbol{q}^*))$ with $\boldsymbol{q}^* \leq \boldsymbol{q}$.

   Since $\boldsymbol{q}^* = \boldsymbol{J}(\boldsymbol{q}^*, \boldsymbol{w}^*(\boldsymbol{q}^*))$,

$$w_{n,m}^*(\boldsymbol{q}) \;=\; \hat{w}_{n,m}(q_n^*, \boldsymbol{q}^*), \quad \forall m \in C_n, \forall n \in \mathcal{N}.$$

   Therefore, $\sum_{m \in C_n} w_{n,m}^*(\boldsymbol{q}) = \sum_{m \in C_n} \hat{w}_{n,m}(q_n^*, \boldsymbol{q}^*) = 1$ for all $n \in \mathcal{N}$. The required result follows.

2. This follows from Corollary 5.3.6.

3. For all $n \in \mathcal{N}$, define $\boldsymbol{w}_n^\dagger$ such that

$$w_{n,m}^\dagger \;\equiv\; \hat{w}_{n,m}(q_n^\dagger, \boldsymbol{q}^\dagger), \quad \forall m \in C_n,$$

and note that $\boldsymbol{w}_n^\dagger$ is feasible since $\sum_{m\in C_n} w_{n,m}^\dagger = 1$. Thus, for each $n \in \mathcal{N}$, the weight vector $\boldsymbol{w}_n^\dagger$ satisfies the normalized rate requirements of all users $m \in C_n$ when the power vector is $\boldsymbol{q}^\dagger$, that is, $(\boldsymbol{q}^\dagger, \boldsymbol{w}^\dagger)$ satisfies (5.5).

Suppose that $\boldsymbol{q}^\dagger$ is not the minimal solution. Then, $\boldsymbol{q}^* \leq \boldsymbol{q}^\dagger$ and $\boldsymbol{q}^* \neq \boldsymbol{q}^\dagger$. Thus, there is a cell $u = \arg\min_{n\in\mathcal{N}} q_n^* \,/\, q_n^\dagger$ with $\alpha_u = q_u^*/q_u^\dagger < 1$. Lemma 5.3.2, together with (5.17), imply

$$\sigma_n(\alpha_u q_u^\dagger, \alpha_u \boldsymbol{q}^\dagger) \;>\; 1.$$

From the fact that $\alpha_u q_u^\dagger = q_u^*$ and $\alpha_u \boldsymbol{q}^\dagger \leq \boldsymbol{q}^*$, and since $\sigma_n(\cdot, \boldsymbol{q})$ is a monotonically increasing function of $\boldsymbol{q}$, we have, $\sigma_n(q_u^*, \boldsymbol{q}^*) > 1$, which contradicts Lemma 5.3.1, since $\boldsymbol{q}^*$ satisfies (5.5). Therefore $\boldsymbol{q}^\dagger = \boldsymbol{q}^*$.

□

We devise two iterative algorithms to solve Problem (5.4). First of these will make use of the standard power control algorithm [106].

5.3.3 Using the Standard Power Control Algorithm

Our first distributed algorithm to solve the minimization problem (5.4) utilizes the result, stated in Lemma 5.3.5, that $\boldsymbol{I}(\boldsymbol{q})$ is standard. Since $\boldsymbol{I}(\boldsymbol{q})$ is a standard interference function, the standard power control algorithm [106] can be applied to solve Problem (5.4). If the power control problem is feasible, then the standard power control algorithm will converge to the minimal solution. The standard power control algorithm is simply the iteration $q_n^{(t+1)} = I_n(\boldsymbol{q}^{(t)})$.

The only difficulty in applying the standard power control algorithm to the present case is that the interference function $\boldsymbol{I}(\boldsymbol{q})$ is not directly measurable. The value of $I_n(\boldsymbol{q})$ is obtained by solving the optimization (5.14). Solving for $I_n(\boldsymbol{q})$ at base station n requires only the knowledge of $K_{n,m}(\boldsymbol{q}) = \rho_{n,m} \,/\, \gamma_{n,m}(\rho_{n,m}, \boldsymbol{q})$. Therefore, this algorithm can be executed locally at each base station with the feedback on achieved SIR from only its users, and hence is distributed.

Lemma 5.3.4 provides an effective means to determine the solution to Problem (5.14). The solution $I_n(\boldsymbol{q})$ can be determined by solving $\sigma_n(t_n, \boldsymbol{q}) = 1$ for t_n using standard methods such as Newton's method and bisection method described in [10]. The initial estimate for the Newton's method can be any $q_n > 0$. Eq. (5.13) of Lemma 5.3.3 provides the bounds for the bisection method, that can be used to bracket the solution.

We describe the algorithm here:

Algorithm 5.3.1.

- ***Initialization:*** *Start with any initial transmit power level $q_n^{(0)}$. Set $k = 0$.*
- ***Iteration** k:*

  1. *Using iterative methods, solve $\sigma_n(\bar{t}_n, \mathbf{q}^{(k)}) = 1$ for $\bar{t}_n$.*
  2. *Set $q_n^{(k+1)} = \bar{t}_n$.*

5.4 Single Iteration Algorithm for Scheme 3

In the preceeding section, we developed an algorithm (Algorithm 5.3.1) to solve Problem (5.4) using the standard power control algorithm [106]. The disadvantage of using Algorithm 5.3.1 is that it involves solving a nonlinear optimization problem (5.14) by iterative methods at each step of the power control iteration.

This motivates us to find an algorithm that eliminates the need to have iterations within iterations. In this section, we develop an algorithm which achieves that by interleaving the steps of power control iteration and weight vector optimization.

We describe the algorithm first.

Algorithm 5.4.1.

- ***Initialization:*** *Start with an initial transmit power level $q_n^{(0)}$. Set $k = 0$.*
- ***Iteration** k:*

1. *Compute a weight vector* $\mathbf{w}_n^{(k)}$*, given the vector of powers for the network* $\mathbf{q}^{(k)}$*, by normalizing the pseudo weight vector* $\hat{\mathbf{w}}_n(q_n^{(k)}, \mathbf{q}^{(k)})$*, (using* (5.10) *and* (5.11)*),*

$$w_{n,m}^{(k)} = \frac{\hat{w}_{n,m}(q_n^{(k)}, \mathbf{q}^{(k)})}{\sigma_n(q_n^{(k)}, \mathbf{q}^{(k)})}, \quad \forall m \in C_n.$$

2. *Use the computed weight vector* $\mathbf{w}_n^{(k)}$ *to compute the target transmit power* $t_{n,m}^{(k)}$ *for the users* $m \in C_n$*:*

$$t_{n,m}^{(k)} = J_{n,m}(\mathbf{q}^{(k)}, w_{n,m}^{(k)}), \quad \forall m \in C_n.$$

3. *Compute the transmit power for the next iteration* $q_n^{(k+1)}$ *depending on the value of* $\sigma_n(q_n^{(k)}, \mathbf{q}^{(k)})$*:*

$$q_n^{(k+1)} = \begin{cases} \min\limits_{m \in C_n} t_{n,m}^{(k)} & \text{if } \sigma_n(q_n^{(k)}, \mathbf{q}^{(k)}) > 1 \\ \max\limits_{m \in C_n} t_{n,m}^{(k)} & \text{otherwise} \end{cases}$$

Each iteration of Algorithm 5.4.1 can be considered as a mapping $\boldsymbol{T}$ from $\boldsymbol{q}^{(k)}$ to $\boldsymbol{q}^{(k+1)}$. The mapping $\boldsymbol{T}$ is continuous and consists of two components $\boldsymbol{w}$ and $\boldsymbol{T}'$: the weight vector update $\boldsymbol{w}^{(k)} = \boldsymbol{w}(\boldsymbol{q}^{(k)})$ and the power update $\boldsymbol{q}^{(k+1)} = \boldsymbol{T}'(\boldsymbol{q}^{(k)}, \boldsymbol{w}^{(k)})$. First, note that the function $\boldsymbol{w}(\boldsymbol{q})$ is *not* the solution, $\boldsymbol{w}^*(\boldsymbol{q})$, to the problem (5.14); it is defined above in step 1) of Algorithm 5.4.1. Second, note that the mapping, $\boldsymbol{T}$, is not standard, and hence the results of [106] do not apply.

5.4.1 Properties of Mapping T

Before proceeding with the convergence analysis of Algorithm 5.4.1, we describe some useful properties of the mapping $\boldsymbol{T}$ that will be useful in the convergence analysis.

When the base stations use $(\boldsymbol{q}, \boldsymbol{w}(\boldsymbol{q}))$ for transmission, the status of a base station $n \in \mathcal{N}$ depends on whether $\sigma_n(q_n, \boldsymbol{q})$ is greater than, less than or equal

to 1, as characterized by the following lemma.

Lemma 5.4.1. *Suppose that the base stations use* $(\mathbf{q}, \mathbf{w}(\mathbf{q}))$ *for transmission. Then the following statements hold.*

1. *If* $\sigma_n(q_n, \mathbf{q}) > 1$*, then the transmit power* q_n *of base station* n *is insufficient to meet the rate requirements of any of its users* $m \in C_n$*. Furthermore,*

$$q_n \;<\; T_n(\mathbf{q}) \tag{5.18}$$

and

$$\sigma_n(T_n(\mathbf{q}), \mathbf{q}) \;\geq\; 1. \tag{5.19}$$

2. *If* $\sigma_n(q_n, \mathbf{q}) = 1$*, then the transmit power* q_n *of base station* n *is exactly sufficient to achieve the rate requirements of all of its users. Furthermore,*

$$q_n \;=\; T_n(\mathbf{q}) \tag{5.20}$$

and

$$\sigma_n(T_n(\mathbf{q}), \mathbf{q}) \;=\; 1. \tag{5.21}$$

3. *If* $\sigma_n(q_n, \mathbf{q}) < 1$*, then the transmit power* q_n *of base station* n *is more than sufficient. Furthermore,*

$$q_n \;>\; T_n(\mathbf{q}) \tag{5.22}$$

and

$$\sigma_n(T_n(\mathbf{q}), \mathbf{q}) \;\leq\; 1. \tag{5.23}$$

Proof.

1. User $m \in C_n$ requires a proportion $\hat{w}_{n,m}(q_n, \boldsymbol{q})$ of subchannels in order to achieve its rate requirement when the transmit power vector is $\boldsymbol{q}$. Since $\sigma_n(q_n, \boldsymbol{q}) > 1$, we have that $w_{n,m}(\boldsymbol{q}) < \hat{w}_{n,m}(q_n, \boldsymbol{q})$ for all $m \in C_n$. Since $J_{n,m}(\cdot, w_{n,m})$ is a monotonically decreasing function of $w_{n,m}$, we have that $J_{n,m}(\boldsymbol{q}, w_{n,m}(\boldsymbol{q})) > J_{n,m}(\boldsymbol{q}, \hat{w}_{n,m}(\boldsymbol{q})) = q_n$ for all $m \in C_n$, i.e., with bandwidth allocation $\boldsymbol{w}_n(\boldsymbol{q})$, the transmit power q_n is insufficient to satisfy the rate requirements of any of the users $m \in C_n$.

   Consequently, $q_n < \min_{m \in C_n} J_{n,m}(\boldsymbol{q}, w_{n,m}(\boldsymbol{q})) = T_n(\boldsymbol{q})$, i.e., (5.18) holds. Since $T_n(\boldsymbol{q})$ is only just sufficient to satisfy the user(s) corresponding to the minimum transmit power level requirement, $\sigma_n(T_n(\boldsymbol{q}), \boldsymbol{q}) \geq 1$, i.e., (5.19) holds.

2. When $\sigma_n(q_n, \boldsymbol{q}) = 1$, we have $w_{n,m}(\boldsymbol{q}) = \hat{w}_{n,m}(q_n, \boldsymbol{q})$ for all $m \in C_n$ and $J_{n,m}(\boldsymbol{q}, w_{n,m}(\boldsymbol{q})) = q_n$ for all $m \in C_n$, i.e., the rate requirement of all users will be met with equality. Furthermore, $q_n = T_n(\boldsymbol{q})$ and hence (5.20) and (5.21) hold.

3. When $\sigma_n(q_n, \boldsymbol{q}) < 1$, we have $w_{n,m}(\boldsymbol{q}) > \hat{w}_{n,m}(q_n, \boldsymbol{q})$ for all $m \in C_n$ and $J_{n,m}(\boldsymbol{q}, w_{n,m}(\boldsymbol{q})) < J_{n,m}(\boldsymbol{q}, \hat{w}_{n,m}(\boldsymbol{q})) = q_n$ for all $m \in C_n$, i.e., with bandwidth allocation $\boldsymbol{w}_n(\boldsymbol{q})$, the transmit power q_n is more than sufficient to meet the rate requirements of all users. Consequently, $q_n > \max_{m \in C_n} J_{n,m}(\boldsymbol{q}, w_{n,m}(\boldsymbol{q})) = T_n(\boldsymbol{q})$, i.e., (5.22) holds. Since $T_n(\boldsymbol{q})$ is sufficient to satisfy each of the users, $\sigma_n(T_n(\boldsymbol{q}), \boldsymbol{q}) \leq 1$, i.e., (5.23) holds.

□

Although $\boldsymbol{T}$ does not satisfy the key monotonicity condition required in Yates framework [106], the following convergence analysis exploits some key inequalities related to the mapping $\boldsymbol{T}$. These properties imply that to any sequence, $(\boldsymbol{q}^{(k)})_{k=1}^{\infty}$, generated by Algorithm 5.4.1, a monotonically non-increasing upper bounding sequence, and a monotonically non-decreasing lower bounding sequence, can always be found with the property that both bounding sequences provably converge to the solution of (5.4).

The following definitions will be used to quantify the bounds on the transmit power levels.

Definition 5.4.1.

$$\alpha(\boldsymbol{q}) \equiv \min_{n\in\mathcal{N}} \frac{q_n}{q_n^*}$$

and

$$\beta(\boldsymbol{q}) \equiv \max_{n\in\mathcal{N}} \frac{q_n}{q_n^*}$$

where $\boldsymbol{q}^*$ is the optimal solution of (5.4).

These definitions immediately imply that

$$\alpha(\boldsymbol{q})q_n^* \leq q_n \leq \beta(\boldsymbol{q})q_n^*, \qquad \forall n \in \mathcal{N}. \tag{5.24}$$

The following lemma and its corollary provide the main properties of $\boldsymbol{T}$ that we will use in the convergence analysis.

Lemma 5.4.2.

1. *If* $\sigma_n(q_n, \mathbf{q}) > 1$, *then*

$$\sigma_n(T_n(\mathbf{q}), \mathbf{q}) \geq 1, \tag{5.25}$$

$$\alpha(\mathbf{q})q_n^* \leq q_n < T_n(\mathbf{q}). \tag{5.26}$$

2. *If* $\sigma_n(q_n, \mathbf{q}) = 1$, *then*

$$\sigma_n(T_n(\mathbf{q}), \mathbf{q}) = 1, \tag{5.27}$$

$$\alpha(\mathbf{q})q_n^* \leq q_n = T_n(\mathbf{q}) \leq \beta(\mathbf{q})q_n^*. \tag{5.28}$$

3. *If* $\sigma_n(q_n, \mathbf{q}) < 1$, *then,*

$$\sigma_n(T_n(\mathbf{q}), \mathbf{q}) \leq 1, \tag{5.29}$$

$$T_n(\mathbf{q}) < q_n \leq \beta(\mathbf{q}) q_n^*. \tag{5.30}$$

4. *If* $\sigma_n(q_n, \mathbf{q}) \geq 1$ *and* $\beta(\mathbf{q}) \geq 1$, *then*

$$T_n(\mathbf{q}) \leq \beta(\mathbf{q}) q_n^*. \tag{5.31}$$

5. *Part 4) also holds with strict inequalities.*

6. *If* $\sigma_n(q_n, \mathbf{q}) \leq 1$ *and* $\alpha(\mathbf{q}) \leq 1$, *then*

$$\alpha(\mathbf{q}) q_n^* \leq T_n(\mathbf{q}). \tag{5.32}$$

7. *Part 6) also holds with strict inequalities.*

8. *If* $\mathbf{q} \geq \mathbf{q}^*$ *and* $\sigma_n(q_n, \mathbf{q}) \leq 1$, *then* $T_n(\mathbf{q}) \geq q_n^*$.

9. *If* $\mathbf{q} \leq \mathbf{q}^*$ *and* $\sigma_n(q_n, \mathbf{q}) \geq 1$, *then* $T_n(\mathbf{q}) \leq q_n^*$.

Proof.

1. This follows from Lemma 5.4.1 (1) and (5.24).

2. This follows from Lemma 5.4.1 (2) and (5.24).

3. This follows from Lemma 5.4.1 (3) and (5.24).

4. Using Lemma 5.3.2 and Theorem 5.3.7(1), we have $\sigma_n(\beta(\boldsymbol{q}) q_n^*, \beta(\boldsymbol{q}) \boldsymbol{q}^*) \leq 1$ since $\beta(\boldsymbol{q}) \geq 1$. Using $\boldsymbol{q} \leq \beta(\boldsymbol{q}) \boldsymbol{q}^*$, this becomes,

$$\sigma_n(\beta(\boldsymbol{q}) q_n^*, \boldsymbol{q}) \leq 1,$$

since $\sigma_n(\cdot, \boldsymbol{q})$ is monotonically increasing in $\boldsymbol{q}$. Since $\sigma_n(q_n, \boldsymbol{q}) \geq 1$, using 1) and 2), we obtain

$$\sigma_n(T_n(\boldsymbol{q}), \boldsymbol{q}) \quad \geq \quad \sigma_n(\beta(\boldsymbol{q})q_n^*, \boldsymbol{q}).$$

Therefore, $T_n(\boldsymbol{q}) \leq \beta(\boldsymbol{q})q_n^*$, since $\sigma_n(t, \cdot)$ is monotonically decreasing in t.

5. This follows from an analogous argument to that in 4), with strict inequalities except for $\boldsymbol{q} \leq \beta(\boldsymbol{q})\boldsymbol{q}^*$.

6. Using Lemma 5.3.2 and Theorem 5.3.7(1), we have $\sigma_n(\alpha(\boldsymbol{q})q_n^*, \alpha(\boldsymbol{q})\boldsymbol{q}^*) \geq 1$ since $\alpha(\boldsymbol{q}) \leq 1$. Using the fact that $\alpha(\boldsymbol{q})\boldsymbol{q}^* \leq \boldsymbol{q}$, it becomes,

$$\sigma_n(\alpha(\boldsymbol{q})q_n^*, \boldsymbol{q}) \quad \geq \quad 1,$$

since $\sigma_n(\cdot, \boldsymbol{q})$ is monotonically increasing in $\boldsymbol{q}$. Since $\sigma_n(q_n, \boldsymbol{q}) \leq 1$, using 2) and 3), we obtain

$$\sigma_n(\alpha(\boldsymbol{q})q_n^*, \boldsymbol{q}) \quad \geq \quad \sigma_n(T_n(\boldsymbol{q}), \boldsymbol{q}).$$

Therefore, $\alpha(\boldsymbol{q})q_n^* \leq T_n(\boldsymbol{q})$, since $\sigma_n(t, \cdot)$ is monotonically decreasing in t.

7. This follows from an analogous argument to that in 6), with strict inequalities except for $\alpha(\boldsymbol{q})\boldsymbol{q}^* \leq \boldsymbol{q}$.

8. Since $\boldsymbol{q} \geq \boldsymbol{q}^*$, using Theorem 5.3.7(1) gives $\sigma_n(q_n^*, \boldsymbol{q}) \geq 1$ since $\sigma_n(\cdot, \boldsymbol{q})$ is monotonically increasing in $\boldsymbol{q}$. Since $\sigma_n(q_n, \boldsymbol{q}) \leq 1$, using 2) and 3), we obtain

$$\sigma_n(T_n(\boldsymbol{q}), \boldsymbol{q}) \quad \leq \quad \sigma_n(q_n^*, \boldsymbol{q}).$$

Therefore, $T_n(\boldsymbol{q}) \geq q_n^*$, since $\sigma_n(t, \cdot)$ is monotonically decreasing in t.

9. Since $\boldsymbol{q} \leq \boldsymbol{q}^*$, using Theorem 5.3.7(1) and the fact that $\sigma_n(\cdot, \boldsymbol{q})$ is monotonically increasing in $\boldsymbol{q}$, we have, $\sigma_n(q_n^*, \boldsymbol{q}) \leq 1$. Since $\sigma_n(q_n, \boldsymbol{q}) \geq 1$, using 1),

we have,

$$\sigma_n(T_n(\boldsymbol{q}), \boldsymbol{q}) \;\; \geq \;\; \sigma_n(q_n^*, \boldsymbol{q}).$$

Therefore, $T_n(\boldsymbol{q}) \leq q_n^*$, since $\sigma_n(t, \cdot)$ is monotonically decreasing in t.

□

Corollary 5.4.3.

1. *If* $\mathbf{q} \geq \mathbf{q}^*$ *then* $T(\mathbf{q}) \geq \mathbf{q}^*$.
2. *If* $\mathbf{q} \leq \mathbf{q}^*$ *then* $T(\mathbf{q}) \leq \mathbf{q}^*$.
3. *If* $\alpha(\mathbf{q}) \leq 1$, *then,* $T(\mathbf{q}) \geq \alpha(\mathbf{q})\mathbf{q}^*$, *or equivalently,* $\alpha(T(\mathbf{q})) \geq \alpha(\mathbf{q})$.
4. *Part 3) also holds with strict inequalities.*
5. *If* $\beta(\mathbf{q}) \geq 1$, *then* $T(\mathbf{q}) \leq \beta(\mathbf{q})\mathbf{q}^*$, *or equivalently,* $\beta(T(\mathbf{q})) \leq \beta(\mathbf{q})$.
6. *Part 5) also holds with strict inequalities.*

5.4.2 Proof of Convergence

We are now in a position to prove the convergence of Algorithm 5.4.1.

Theorem 5.4.4. *If (5.5) is feasible, then, Algorithm 5.4.1 converges to the solution of Problem (5.4).*

Proof. Consider the following cases.

1. If $1 \leq \alpha(\boldsymbol{q}) \leq \beta(\boldsymbol{q})$, then $\boldsymbol{q}^* \leq \alpha(\boldsymbol{q})\boldsymbol{q}^* \leq \boldsymbol{q} \leq \beta(\boldsymbol{q})\boldsymbol{q}^*$. By Corollary 5.4.3(5), we have $T(\boldsymbol{q}) \leq \beta(\boldsymbol{q})\boldsymbol{q}^*$, and by Corollary 5.4.3(1), we have $T(\boldsymbol{q}) \geq \boldsymbol{q}^*$. It follows that

$$1 \;\; \leq \;\; \alpha(T(\boldsymbol{q})) \;\leq\; \beta(T(\boldsymbol{q})) \;\leq\; \beta(\boldsymbol{q}).$$

Applying this argument inductively to the sequence $\boldsymbol{q}^{(k)} \equiv T^k(\boldsymbol{q})$, we have:

$$\begin{aligned} \boldsymbol{q}^* &\leq \boldsymbol{q}^{(k+1)} \leq \beta(\boldsymbol{q}^{(k)})\boldsymbol{q}^* \text{ and} \\ 1 &\leq \alpha(\boldsymbol{q}^{(k+1)}) \leq \beta(\boldsymbol{q}^{(k+1)}) \leq \beta(\boldsymbol{q}^{(k)}), \quad \forall k. \end{aligned}$$

and thus $\beta(\boldsymbol{q}^{(k)}) \downarrow \beta^* \geq 1$ as $k \to \infty$. Note also that $\boldsymbol{q}^* \leq \boldsymbol{q}^{(k)} \leq \beta(\boldsymbol{q}^{(1)})\boldsymbol{q}^*$ for all k and hence $(\boldsymbol{q}^{(k)})_{k=1}^{\infty}$ is bounded. Let $(\boldsymbol{q}^{(k_i)})_{i=1}^{\infty}$ be a subsequence converging to an accumulation point $\boldsymbol{q}^\dagger$. By continuity of β,

$$\begin{aligned} \beta(\boldsymbol{q}^{(k_i)}) &\to \beta(\boldsymbol{q}^\dagger) = \beta^* \text{ as } i \to \infty \\ \beta(T(\boldsymbol{q}^{(k_i)})) &\to \beta(T(\boldsymbol{q}^\dagger)) = \beta^* \text{ as } i \to \infty \end{aligned} \tag{5.33}$$

By the continuity of $\boldsymbol{T}$, $\boldsymbol{q}^\dagger$ must satisfy

$$\boldsymbol{q}^* \leq T(\boldsymbol{q}^\dagger) \leq \beta(T(\boldsymbol{q}^\dagger))\boldsymbol{q}^* = \beta^*\boldsymbol{q}^* \tag{5.34}$$

If $\beta^* > 1$, then by Corollary 5.4.3(6), $\beta(T(\boldsymbol{q}^\dagger)) < \beta^*$ which is a contradiction. Hence, $\beta^* = 1$ and $\beta(\boldsymbol{q}^{(k)}) \downarrow 1$ as $k \to \infty$. Therefore, $\boldsymbol{q}^{(k)} \to \boldsymbol{q}^*$ from (5.34).

2. If $\alpha(\boldsymbol{q}) \leq \beta(\boldsymbol{q}) \leq 1$, then, $\boldsymbol{q}^{(k)} \to \boldsymbol{q}^*$ follows from an analogous argument to that used in case 1, but with α's replacing β's, and using Corollary 5.4.3, parts 2, 3 and 4, to obtain an increasing sequence of $\alpha(\boldsymbol{q}^{(k)})$'s, in place of a decreasing sequence of $\beta(\boldsymbol{q}^{(k)})$'s.

3. If $\alpha(\boldsymbol{q}^{(k)}) < 1 < \beta(\boldsymbol{q}^{(k)})$, $\forall k$, then parts 4 and 6 of Corollary 5.4.3 imply that $(\alpha(\boldsymbol{q}^{(k)}))_{k=1}^{\infty}$ is a strictly increasing sequence and bounded above by 1, and $(\beta(\boldsymbol{q}^{(k)}))_{k=1}^{\infty}$ is a strictly decreasing sequence and bounded below by 1. Let the limits be α^* and β^* respectively. Then, $\alpha^* \leq 1 \leq \beta^*$. Note also $\alpha(\boldsymbol{q}^{(1)})\boldsymbol{q}^* \leq \boldsymbol{q}^{(k)} \leq \beta(\boldsymbol{q}^{(1)})\boldsymbol{q}^*, \forall k$. Therefore, the sequence $(\boldsymbol{q}^{(k)})_{k=1}^{\infty}$ has accumulation points. Let $(\boldsymbol{q}^{(k_i)})_{i=1}^{\infty}$ be a subsequence of points converging to an accumulation point $\boldsymbol{q}^\dagger$. By the continuity of α,

$$\alpha(\boldsymbol{q}^{(k_i)}) \to \alpha(\boldsymbol{q}^\dagger) = \alpha^* \text{ as } i \to \infty$$

$$\alpha(T(\boldsymbol{q}^{(k_i)})) \;\; \rightarrow \;\; \alpha(T(\boldsymbol{q}^{\dagger})) = \alpha^* \text{ as } i \rightarrow \infty \tag{5.35}$$

and similarly, by the continuity of β, (5.33) holds. Therefore,

$$\begin{aligned} \alpha(\boldsymbol{q}^{\dagger}) &= \alpha(T(\boldsymbol{q}^{\dagger})) = \alpha^* \leq 1, \\ \beta(\boldsymbol{q}^{\dagger}) &= \beta(T(\boldsymbol{q}^{\dagger})) = \beta^* \geq 1 \end{aligned}$$

and $\alpha^*\boldsymbol{q}^* \leq T(\boldsymbol{q}^{\dagger}) \leq \beta^*\boldsymbol{q}^*$. If $\alpha^* < 1$, then by Corollary 5.4.3(4), $\alpha(T(\boldsymbol{q}^{\dagger})) > \alpha^*$, which is a contradiction. Hence $\alpha^* = 1$. Similarly, if $\beta^* > 1$, then by Corollary 5.4.3(6) $\beta(T(\boldsymbol{q}^{\dagger})) < \beta^*$, which is a contradiction. Hence $\beta^* = 1$. We conclude that $\boldsymbol{q}^{\dagger} = \boldsymbol{q}^*$.

4. There exists a $\kappa \in \mathbb{Z}_+$ such that $\alpha(\boldsymbol{q}^{(k)}) < 1 < \beta(\boldsymbol{q}^{(k)})$ for $k = 1, \ldots, \kappa-1$ and, $\alpha(\boldsymbol{q}^{(\kappa)}) \geq 1$ or $\beta(\boldsymbol{q}^{(\kappa)}) \leq 1$. Then, at iteration $k = \kappa$, it falls under case 2) if $\beta(\boldsymbol{q}^{(\kappa)}) \leq 1$ or case 1) if $\alpha(\boldsymbol{q}^{(\kappa)}) \geq 1$.

□

5.5 Numerical Studies

In this section, we numerically study the features of the three schemes that we have developed to solve Problem (4.5), by simulations.

The simulation environment that we used for this evaluation is as follows. We consider a cellular network that spans an area of 3km $\times$ 3km. This area is divided into 9 square cells of size 1km $\times$ 1km. The base stations are located at the center of their respective cells (refer to Figure 5.3). There are 250 users, uniformly randomly distributed in the network. Each user communicates only to the base station of the cell it is located. We use free-space propagation for distances up to 500 m, and a log distance path loss, with exponent 3, beyond that (relative to the reference distance of 500 m). For all distances we include a log-normal shadowing with a mean of 0 dB and a standard deviation of 8 dB. In this context, a network realization refers to a random realization of user locations and shadowing.

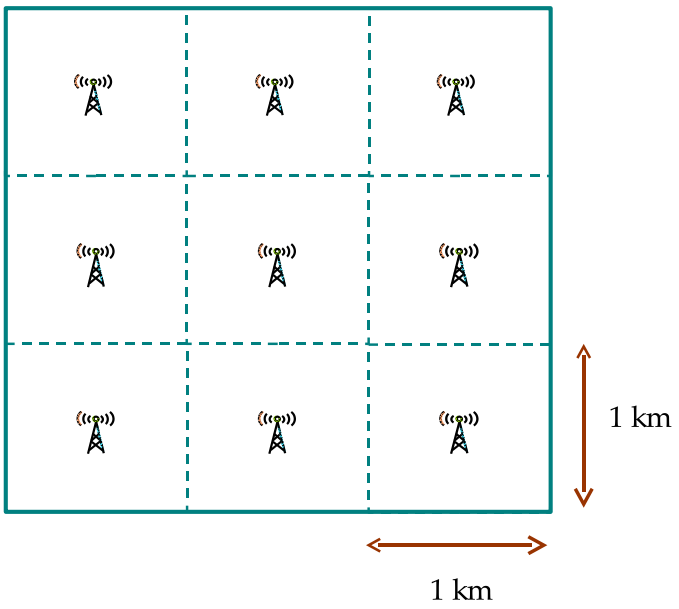


Figure 5.3: Simulation Scenario

The normalized rate target for each user is selected from a set of normalized rate targets available in the network. Firstly, we consider the special case that only one normalized rate target is available for the users to choose from, i.e., every user in the network has the same normalized rate target. We will then study the case in which there are four normalized rate targets available in the network for the users to select from.

The noise power spectral density at each receiver is 10^{-19} W/Hz. The achieved normalized rate of a link is computed using Shannon formula, i.e., the spectral efficiency of a link with a SIR of γ is taken be $f(\gamma) = \log_2(1+\gamma)$ (spectral efficiency f is characterized in Section 4.2).

5.5.1 Equal Rates

In this section, we consider the case in which every user has the same normalized rate target, say c bits/sec/Hz. By tuning the value of c, the normalized rate targets of the users can be varied. We study various aspects of the schemes below.

Convergence and the Initial Condition

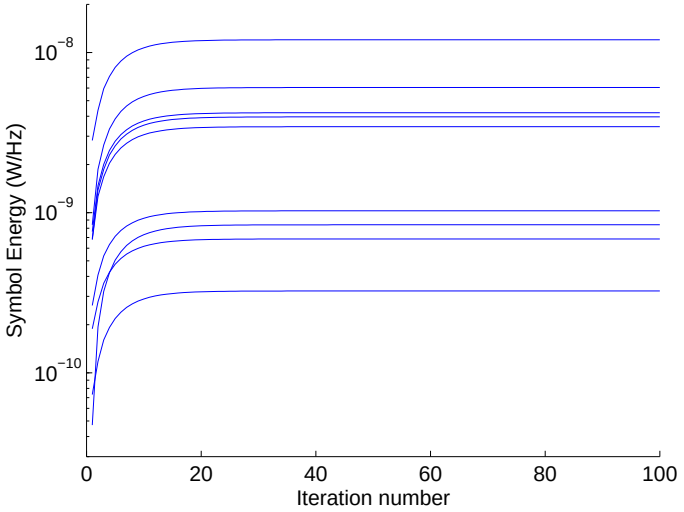
Recall that all three schemes will converge to their respective optimal solutions regardless of the initial conditions. In order to illustrate this, we studied the convergence of the schemes for various initial conditions. For a particular network realization and an arbitrarily chosen value of c, Figure 5.4 plots the convergence of the symbol energy in each cell for Scheme 1 as a function of the iteration number for two different initial conditions. The corresponding results for Schemes 2 and 3 are not presented here to avoid two much repetition.

Distribution of Transmit Power Densities

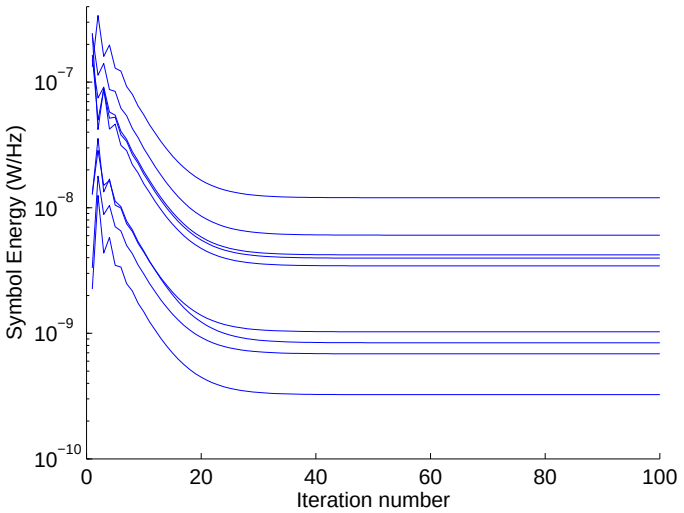
Of the three schemes we considered, Schemes 1 and 2 allow the transmit power densities to be selected on a user-by-user basis. This is in contrast to Scheme 3 which constrains each cell to use a uniform power density across all users in the cell. We study how the power densities vary within each cell for Schemes 1 and 2 compared to Scheme 3.

Figure 5.5 plots the transmit power densities of the users for the three schemes, for a particular network realization and an arbitrary c. Although both Schemes 1 and 2 exhibit variations in the power densities among the same cell users, the variations with Scheme 1 is small as compared to Scheme 2. In fact, the distribution of the power densities of Scheme 1 is actually very similar to that of Scheme 3, in contrast to Scheme 2.

The large variation in the power densities among the same cell users with Scheme 2 can be explained as follows. As Scheme 2 uses a fixed bandwidth allocation, this scheme adapts the power densities to combat the interference and achieve the target SIR. A user with good channel conditions requires a small amount of power for the allocated bandwidth to achieve the desired SIR as opposed to a user with bad channel conditions, leading to a large range in the transmit power densities.



(a) Initial power densities are set to 0



(b) Initial power densities are chosen randomly with a normal distribution (with mean of 0 and standard deviation of 10^{-7} and taking the absolute values)

Figure 5.4: Equal rates: convergence of symbol energy in each cell as a function of the iteration number for Scheme 1

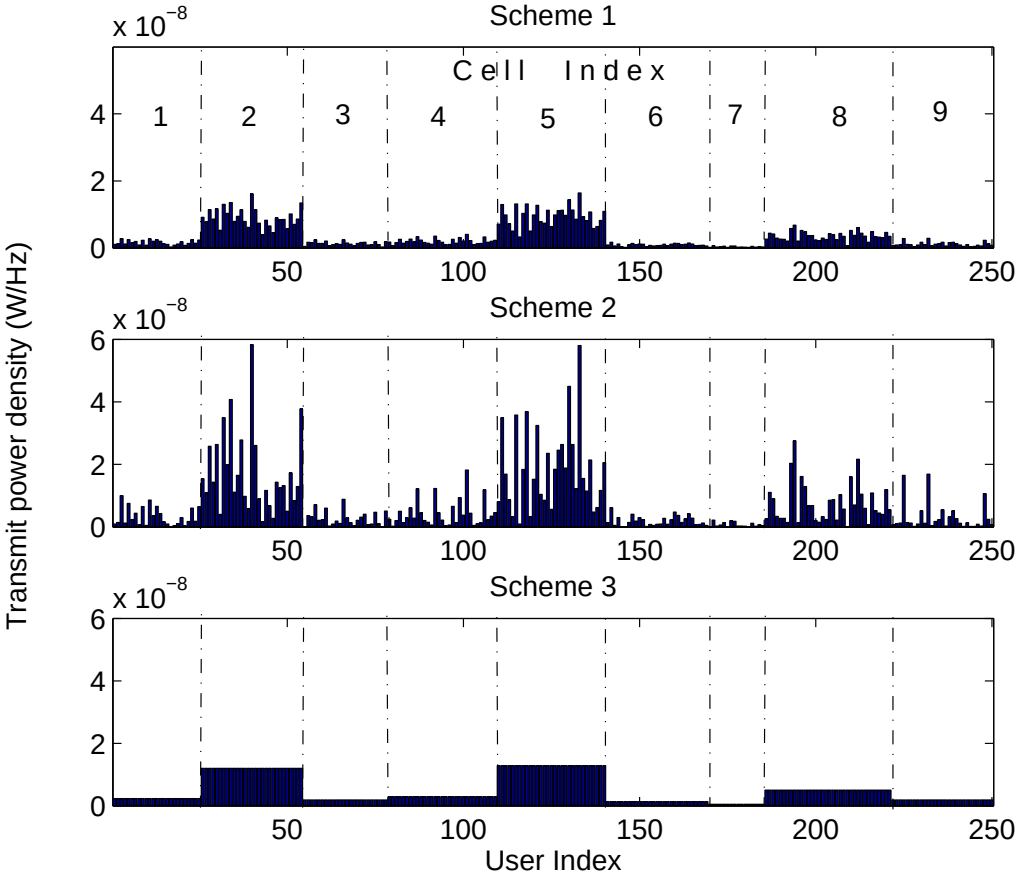


Figure 5.5: Equal rates: transmit power densities of the users

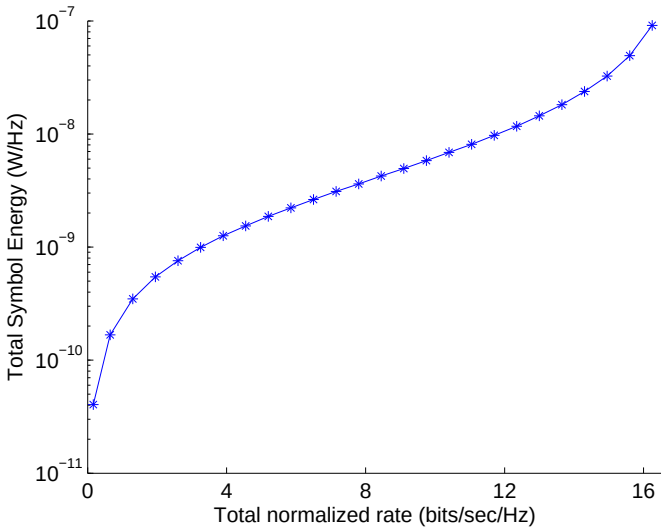


Figure 5.6: Equal rates: total symbol energy consumption against total normalized rate for Scheme 1 for a particular network realization

Qualitative Behaviours of the Schemes

We examine the total energy consumption and the rate of convergence of the schemes as a function of the total normalized rate targets next. Figure 5.6 plots the total symbol energy consumption for Scheme 1, with varying normalized rate targets, for a particular network realization. The normalized rate targets were varied by tuning the scaling factor c. There is a limit to how high we can scale up c; there is an energy explosion that occurs as the capacity limit is reached for that network realization.

Figure 5.7 plots the number of iterations it takes for Scheme 1 to converge to within 0.1%, with varying normalized rate targets. The number of iterations it takes to converge increases with the normalized rate requirements, as expected.

Again, we avoid presenting the results for the other two schemes to avoid repetition, as all three schemes exhibit the same qualitative behaviour in terms of energy consumption as well as rate of convergence. However, the comparison of the three schemes in terms of average symbol energy consumption is presented at the end of this subsection.

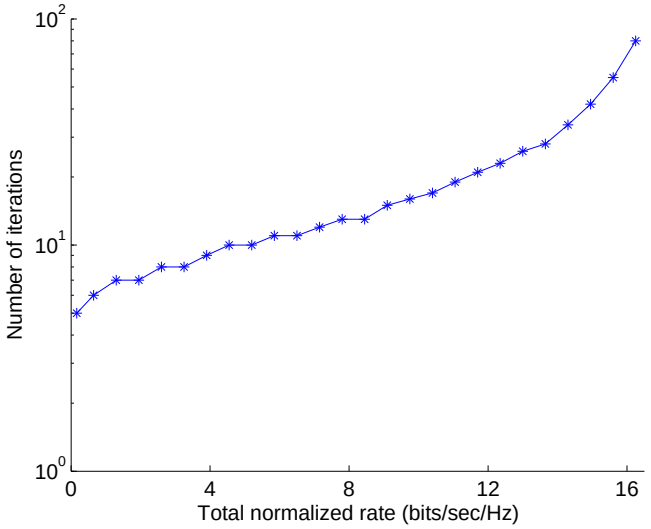


Figure 5.7: Equal rates: iteration count to within 0.1 % of convergence against total normalized rate for Scheme 1 for a particular network realization

Rates of Convergence of the Algorithms That Solve Scheme 3

Recall that we have developed two algorithms to solve the uniform transmit PSD allocation scheme (Scheme 3). While Algorithm 5.3.1 is based on the standard power control algorithm, where the optimal weight vector is computed for the given interference at each step of the power control algorithm, Algorithm 5.4.1 combines the power update and the weight optimization into a single iteration. Will not solving for the optimal weight vector at each step of the power control step affect the rate of convergence of Algorithm 5.4.1, i.e., will it take longer to converge?

Figure 5.8 plots the number of iterations each algorithm takes to converge within 0.1% of the solution with varying normalized rates. Algorithm 5.4.1 requires only a few more iterations than Algorithm 5.3.1, despite using a suboptimal weight allocation at each step. The benefit of computing the optimal weight vector at each step of the algorithm is marginal. This could be explained as follows. The optimal weight vector is a function of the current interference. The interference will be changing from one iteration to the next. Consequently, the optimal weight vector for one iteration is likely to be very different from the optimal weight vector for the previous iteration. Hence, computing the optimal weight

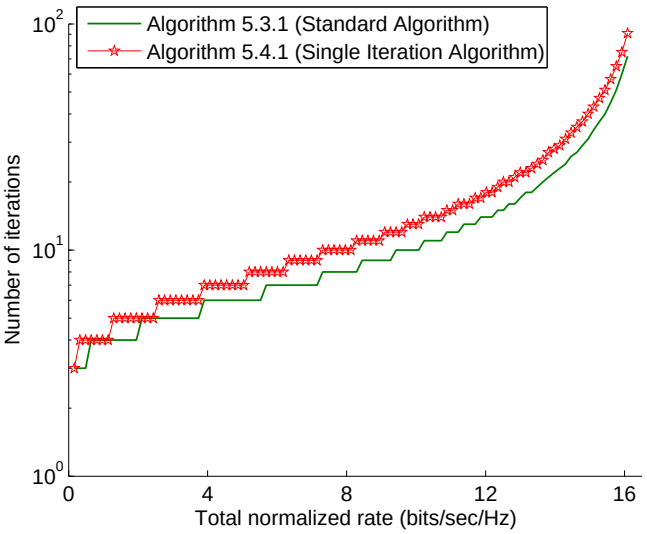


Figure 5.8: Number of iterations to converge to within 0.1% of the solution against total normalized rate (bits/sec/Hz) for the two algorithms that solve Scheme 3

vector may not improve the rate of convergence of the algorithm significantly.

Average Performance of the Schemes

In this section, we compare the three schemes that we have developed. In particular, we quantify the loss in capacity due to the reduction in the available degrees of freedom for optimization. We use the average symbol energy consumption as the performance metric.

For each network realization, the total symbol energies can be computed for a set of values of c, over the entire range of feasible values. The simplest way to compute the average total symbol energy for a given total normalized rate is to average the total symbol energies over many network realizations for that total normalized rate. As different network realizations explode at different total normalized rates, this method of averaging does not give an accurate picture of the performance of the schemes.

A better method is to average the total normalized rates over many network realizations for a given set of values of total symbol energies. However, it will be computationally intensive to solve exactly for the normalized rate for a given symbol energy as the natural numerical procedure is to solve for symbol energies

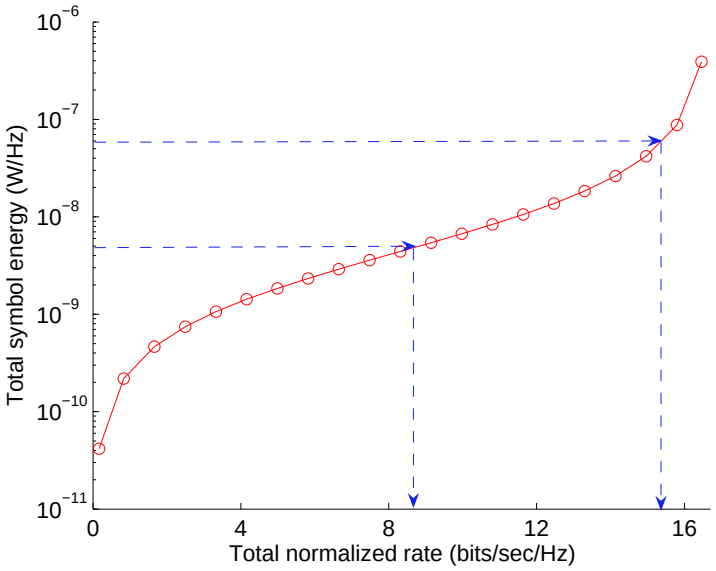


Figure 5.9: Interpolation explained: total symbol energy vs. total normalized rate using Scheme 3 for a particular realization of the network. For a given value of total symbol energy, the corresponding total normalized rate can be computed by linear interpolation.

for given values of normalized rates.

For this reason, we use interpolation to obtain the average total normalized rates. The interpolation is done as follows. For each network realization, the total symbol energies are computed for a set of values of c. Then, for a fixed set of total symbol energies, corresponding total normalized rates are computed by linear interpolation as demonstrated in Figure 5.9. The average total normalized rate for a given total symbol energy is the average of such computed total normalized rates over many network realizations.

Figure 5.10 plots the average total symbol energy versus average total normalized rates for the three schemes averaged over 100 realizations. Notice that all three schemes suffer an energy explosion, as their average capacity limits are reached.

From Figure 5.10, it is clear that Scheme 1 uses the least amount of energy for a given total normalized rate. This is not surprising since Scheme 1 solves an optimization problem with more degrees of freedom. The loss in performance in terms of capacity for a given total symbol energy for Schemes 2 and 3 is small

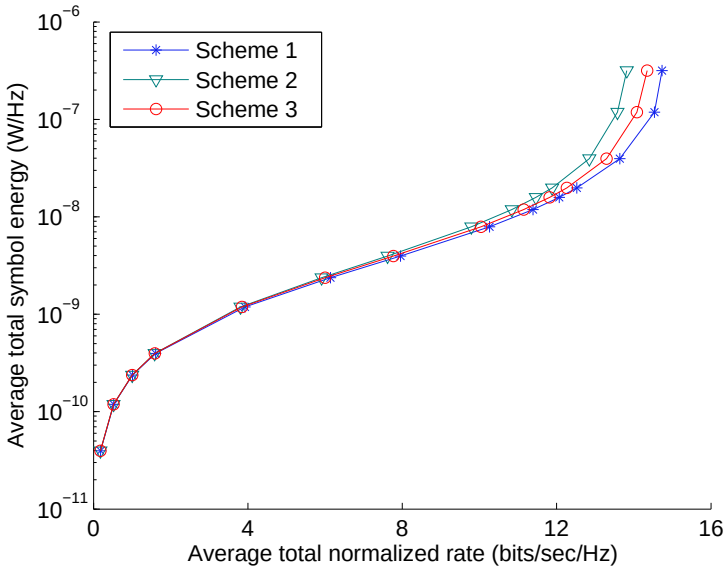


Figure 5.10: Equal rates: average total symbol energy vs. average total normalized rate for the network

compared to Scheme 1 which has all the degrees of freedom available. Of the two reduced complexity schemes, Scheme 3 marginally outperforms Scheme 2. The capacity penalty for Scheme 2 is around twice that of Scheme 3.

5.5.2 Multiple Rates

In this subsection, we consider the case in which there are four normalized rate targets available for the users to select from. Each user randomly selects its normalized rate target from a set of normalized rates $\{c, 2c, 3c, 4c\}$, where c is a scaling factor which is used to vary the normalized rate targets of the users.

The qualitative behaviours of the schemes under this setting are the same as that of the equal rates case. Therefore, we only present a selected set of results for the multiple rates case.

Distribution of Transmit Power Densities

Figure 5.11 shows the transmit power densities of the users for various schemes, for a particular network realization and an arbitrary c. The variations in the power densities is small with Scheme 1 as compared to Scheme 2. Furthermore,

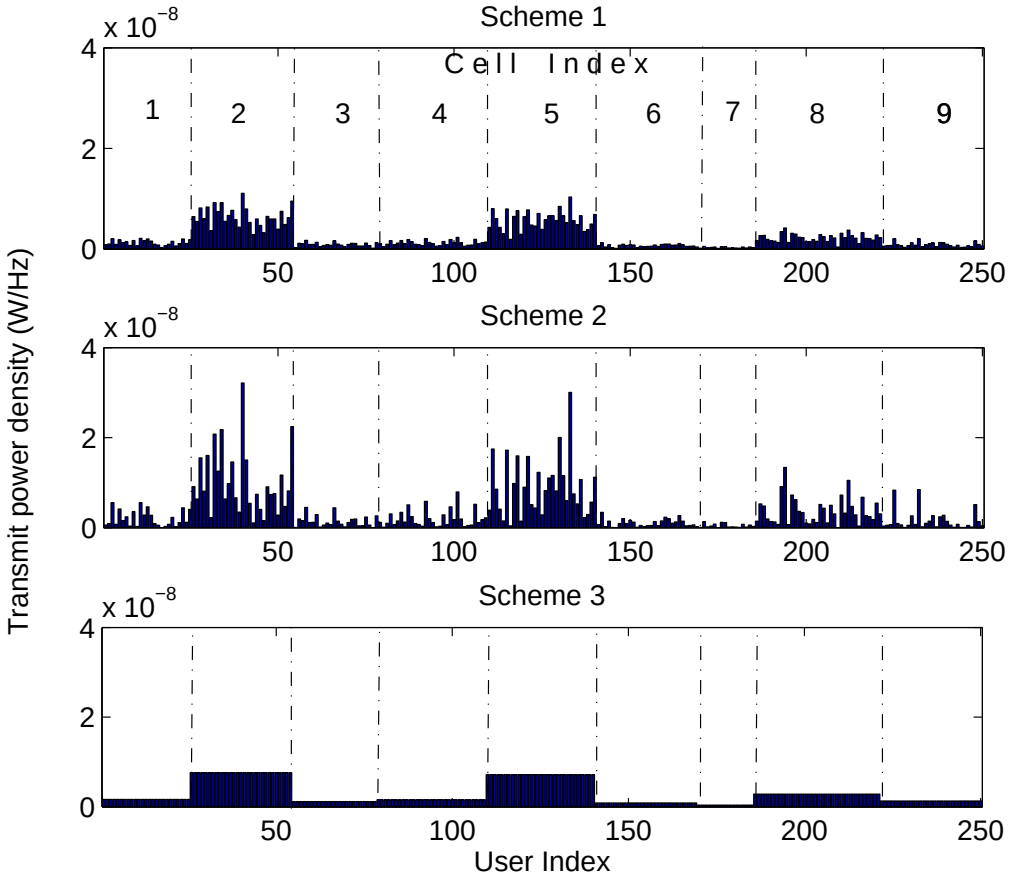


Figure 5.11: Multiple rates: transmit power densities of the users

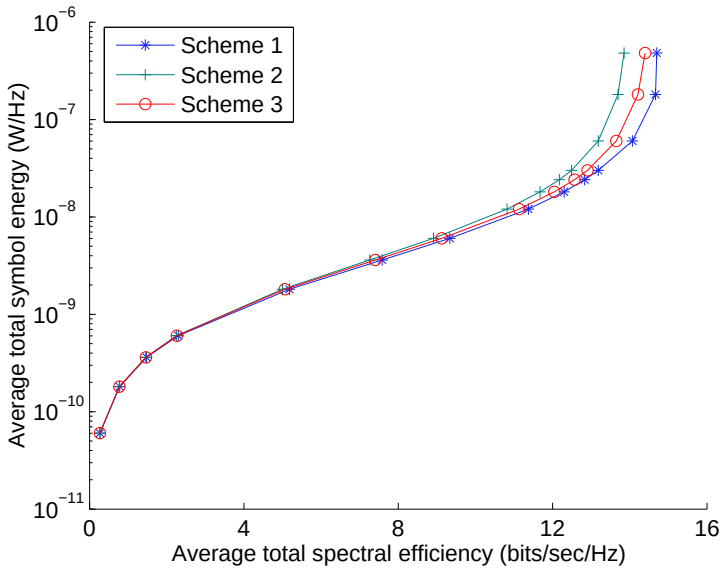


Figure 5.12: Multiple rates: average total symbol energy vs. average total normalized rate for the network

the distribution of the power densities of Scheme 1 looks closer to that of Scheme 3 than that of Scheme 2.

Average Performance of the Schemes

We now compare the performance of the three schemes under the multiple rates regime using the average symbol energy consumption as the performance metric.

The plot of average total symbol energy vs. average total normalized rate in Figure 5.12 is obtained using the procedure described in the corresponding subsection of Section 5.5.1. Schemes 1 and 2 have the best and the worst performances respectively as in the equal rates case. Furthermore, the capacity penalty for Scheme 3 is around a half that suffered by Scheme 2.

5.6 Schemes at a Glance

We have developed three resource allocation schemes for OFDM cellular network under flat fading. Scheme 1 dynamically allocates both subchannels and power spectral densities to each user, and in general the transmit power spectrum of each base station will not be flat. Scheme 2 allocates subchannels in

a static manner, and then dynamically allocates power spectral densities to the users. Scheme 3 allocates subchannels in a dynamic manner, and allows the base stations to adjust their transmit power dynamically as well, under the constraint that the transmit spectrum is flat.

Scheme 2 incorporates the traditional approach to bandwidth allocation in that the modulation scheme for each user is fixed in advance, based on the rate requirements. It tries to ensure that the required SIR's of the users are achieved. In this approach, a user close to a base station will receive much less power than a user near the cell boundary. As a result, there will be a large dynamic range in the possible interference levels that a user will receive from other base stations, since these base stations are also controlling their transmit power levels to accommodate their users. To obtain the interference averaging, it is necessary to employ a frequency hopping strategy, and one that is coordinated between cell sites. This is true in the Flarion system, with the Latin square hopping patterns designed to provide the requisite interference averaging (see [96]).

In Scheme 1, and also Scheme 3, a user far away from the base station can receive more bandwidth (subchannels), and this is available also to a user experiencing a large amount of interference. Indirectly, allocating more bandwidth will also allocate the user more received power, for the same fixed base station transmit power, but there is an additional degree of flexibility in that the modulation scheme can adapt.

With dynamic bandwidth allocation, the users close to the base station will be allocated a small number of subchannels. The base station can use a larger constellation size and hence pack more information into the symbols for these users. On the other hand, the users further away from the base station will be allocated a larger number of subchannels, and hence a higher rate in symbols per second. In this case, only a small constellation size for the symbols is required to achieve the same bit rate. The downside of this approach is the increased complexity of the coding and adaptive modulation required.

The numerical studies show that both Scheme 2 and Scheme 3, provide solutions that are nearly as good as that of Scheme 1. Of the two, Scheme 3 has

marginally better performance in terms of power levels or capacity.

All these schemes takes account of log-normal shadowing and path loss, but not frequency selective multipath fading. The handling of the multipath frequency selective fading effects is the topic of Chapter 6. For the proposed refinement, Scheme 3 which uses a uniform transmit PSD across the users in each cell is chosen as the method of choice. The reason for this choice will be explained in Chapter 6.

These schemes were developed for a flat fading environment, and as a consequence, it was not necessary to implement a frequency hopping strategy when using Scheme 3. The interference averaging requirement automatically is satisfied since the transmit power spectra of the base stations is flat. However, when Scheme 3 is used for the higher layer resource allocation as part of the two-layer approach in solving the resource allocation under frequency selective fading, a frequency hopping strategy may still be useful in conjunction with Scheme 3 to obtain frequency diversity.

5.7 Uniform PSD Allocation Problem with Discrete Weights

The uniform PSD resource allocation problem with continuous weights was considered in Section 5.3. This section studies the corresponding problem with discrete weights. This corresponds to the following optimization problem:

$$\min_{\boldsymbol{q},\boldsymbol{w}} \sum_{n\in\mathcal{N}} q_n \tag{5.36}$$

such that for all n,

$$\begin{aligned} w_{n,m}\, f\left(\gamma_{n,m}(q_n, \boldsymbol{q})\right) &\geq c_{n,m}^{tar},\ \forall m \in C_n, \\ w_{n,m} &\in \mathbf{\Pi},\ \forall\, m \in C_n, \\ \sum_{m \in C_n} w_{n,m} &= 1, \\ q_n &> 0, \end{aligned}$$

where the set $\mathbf{\Pi}$ is given by (4.36). Problem (5.36) is a mixed integer, nonlinear optimization problem where the variables q_n's are continuous and $w_{n,m}$'s are discrete.

5.7.1 Continuous Relaxation

A suboptimal solution to Problem (5.36) could be found by using an approach similar to that of Section 4.6.1. This involves first solving (5.36) in the continuous domain by dropping the discrete constraints (i.e., solving Problem (5.4) by using either Algorithm 5.3.1 or 5.4.1) and then rounding the continuous weights to obtain a solution. As noted in Section 4.6.1, this approach will not necessarily lead to a good discrete solution and may only be regarded as a heuristic.

5.7.2 Exact Solution

We will now solve Problem (5.36) exactly using an iterative algorithm. By using an approach similar to that of Algorithm 5.3.1, we show that Yates framework can be used to solve the problem. At each step of the power control iteration, each cell will solve a problem of finding an optimal weight vector that minimizes the transmit power of the cell for the given interference. This single cell subproblem is combinatorial.

Subproblem for Cell n

The single cell subproblem for cell n corresponds to finding a weight vector that minimizes the transmit power for the cell, when the interference is generated by

the power vector $\boldsymbol{q}$:

$$\min_{q_n, \boldsymbol{w}_n \in \boldsymbol{\Pi}_n} \; q_n \tag{5.37}$$

subject to

$$\begin{aligned} w_{n,m}\, f\left(\frac{q_n}{K_{n,m}(\boldsymbol{q})}\right) &\geq c_{n,m}^{tar},\ \forall m \in C_n, \\ q_n &> 0, \end{aligned}$$

where $\boldsymbol{\Pi}_n$ is a set of available weights for base station n, given by (4.38). Note here that q_n is the maximum of the transmit powers needed for each user in cell n. Furthermore, Problem (5.37) may have multiple weight vectors solving it.

We rewrite (5.37) to obtain:

$$\min_{\boldsymbol{w}_n \in \boldsymbol{\Pi}_n} \; \max_{m \in C_n} \; J_{n,m}(\boldsymbol{q}, w_{n,m}) \tag{5.38}$$

where $J_{n,m}(\boldsymbol{q}, w_{n,m})$ is the transmit power requirement of user m when the weight is $w_{n,m}$ and the interference is generated by a vector of powers $\boldsymbol{q}$, and is defined in (5.6). We will now show how Problem (5.38) can be solved by making use of the solution of its continuous counterpart (5.14) which we have solved earlier. Considering Problem (5.14), let $\tilde{\boldsymbol{w}}_n$ be the unique weight vector that solves (5.14) and $\tilde{q}_n$ be the corresponding power level. If $q_n^{\dagger}$ is the minimum to (5.38), then, $q_n^{\dagger} \geq \tilde{q}_n$. This follows from the fact that the continuous feasible solution space $\boldsymbol{\Omega}_n$ is a superset of the integral feasible solution space $\boldsymbol{\Pi}_n$, i.e., $\boldsymbol{\Pi}_n \subseteq \boldsymbol{\Omega}_n$.

We define a vector $\bar{\boldsymbol{\eta}}_n = (\bar{\eta}_{n,m})_{m \in C_n}$ by,

$$\bar{\eta}_{n,m} \;=\; \lceil \tilde{\eta}_{n,m} \rceil, \quad \forall m \in C_n, \tag{5.39}$$

where $\tilde{\eta}_{n,m} = N_c\, \tilde{w}_{n,m}$. Since $\bar{\eta}_{n,m} \geq \tilde{\eta}_{n,m}$, we have that $\sum_{m \in C_n} \bar{\eta}_{n,m} \geq N_c$. Now define a corresponding $\bar{\boldsymbol{w}}_n = (\bar{w}_{n,m})_{m \in C_n}$:

$$\bar{w}_{n,m} \;=\; \frac{\bar{\eta}_{n,m}}{N_c}, \quad \forall m \in C_n. \tag{5.40}$$

Note that if $\sum_{m\in C_n} \bar{\eta}_{n,m} > N_c$, then, $\bar{\boldsymbol{w}}_n$ is infeasible (i.e., $\sum_{m\in C_n} \bar{w}_{n,m} > 1$).

We denote the transmit power requirement of user m by $q_{n,m}(\eta_{n,m})$, when the number of subchannels allocated to that user is $\eta_{n,m}$, i.e., $w_{n,m} = \eta_{n,m}/N_c$. Then,

$$q_{n,m}(\eta_{n,m}) \;=\; J_{n,m}\left(\boldsymbol{q}, \frac{\eta_{n,m}}{N_c}\right).$$

Clearly, $q_{n,m}(\eta_{n,m})$ is monotonically decreasing in $\eta_{n,m}$ since $J_{n,m}(\cdot, w_{n,m})$ is monotonically decreasing in $w_{n,m}$. Furthermore, $q_{n,m}(\bar{\eta}_{n,m}) \leq \tilde{q}_n \leq q_n^\dagger$, where the first inequality stems from the fact that $q_{n,m}(\cdot)$ is a monotonically decreasing function, $\bar{\eta}_{n,m} \geq \tilde{\eta}_{n,m}$ and $q_{n,m}(\tilde{\eta}_{n,m}) = \tilde{q}_n$.

If $\sum_{m\in C_n} \bar{\eta}_{n,m} = N_c$, then, Problem (5.14) has a discrete solution and $q_n^\dagger = \tilde{q}_n$. Then the solution to (5.14) is also the solution to (5.38).

Turning our attention to the typical case that $\sum_{m\in C_n} \bar{\eta}_{n,m} > N_c$, we note that $q_{n,m}(\bar{\eta}_{n,m}) \leq q_n^\dagger$, $\forall m \in C_n$. Thus, we can apply the following lemma and reduce the excess in the number of subchannels allocated.

Lemma 5.7.1. *For any $\boldsymbol{\eta}_n = (\eta_{n,m})_{m\in C_n}$, if*

- $q_{n,m}(\eta_{n,m}) \leq q_n^\dagger$, $\forall m \in C_n$, *and*
- $\sum_{m\in C_n} \eta_{n,m} > N_c$,

then,

$$\min_{m\in C_n} q_{n,m}(\eta_{n,m} - 1) \;\leq\; q_n^\dagger.$$

Proof. Let $\boldsymbol{\eta}_n^\dagger$ be a subchannel allocation (not necessarily unique) that solves Problem (5.38). Since $\sum_{m\in C_n} \eta_{n,m} > N_c$ and $\sum_{m\in C_n} \eta_{n,m}^\dagger = N_c$, there exists a user m' such that $\eta_{n,m'} > \eta_{n,m'}^\dagger$, or equivalently, $(\eta_{n,m'} - 1) \geq \eta_{n,m'}^\dagger$. Since $q_{n,m}(\cdot)$ is monotonically decreasing, we have $q_{n,m'}(\eta_{n,m'} - 1) \leq q_{n,m'}(\eta_{n,m'}^\dagger) \leq q_n^\dagger$ where the second inequality follows from the fact that $q_n^\dagger = \max_{m\in C_n} q_{n,m}(\eta_{n,m}^\dagger)$. The required result follows. □

We now show how to construct a discrete solution to Problem (5.38) by successively applying Lemma 5.7.1, starting with the allocation in $\bar{\boldsymbol{\eta}}_n$ and terminating

after a finite number of steps in such a solution.

We start with $\boldsymbol{\eta}_n^{(0)} = \bar{\boldsymbol{\eta}}_n$ and note that this allocation involves removing $r_n = \sum_{m \in C_n} \bar{\eta}_{n,m} - N_c$ excess subchannels. We do this iteratively.

Suppose $\boldsymbol{\eta}_n^{(k)}$ has $r_n - k > 0$ excess subchannels, and satisfies $q_{n,m}(\eta_{n,m}^{(k)}) \leq q_n^{\dagger}$, for all $m \in C_n$. Then Lemma 5.7.1 applies. In particular, removing a subchannel from user m that satisfies

$$m = \arg \min_{m \in C_n} \; q_{n,m}(\eta_{n,m}^{(k)} - 1)$$

results in a new subchannel allocation $\boldsymbol{\eta}_n^{(k+1)}$ that has $r_n - (k+1)$ excess subchannels, and satisfies $q_{n,m}(\eta_{n,m}^{(k+1)}) \leq q_n^{\dagger}$, $\forall m \in C_n$. By induction, we obtain after r_n steps that there are no excess subchannels, so we must be at a solution to Problem (5.38). Note that $r_n < L_n$ where L_n is the number of users in cell n.

The algorithm for solving (5.38) is summarized below.

Algorithm 5.7.1.

- ***Initialization:*** *Solve the continuous relaxation of* (5.37) *which is* (5.14). *Set* $\boldsymbol{\eta}_n^{(0)} = \bar{\boldsymbol{\eta}}_n$.
- *Remove r_n excess subchannels from users, one at a time.*

for $k = 1$ *to* r_n ***do***
$\boldsymbol{\eta}_n^{(k+1)} = \boldsymbol{\eta}_n^{(k)}$
$m' \leftarrow \arg \min_{m \in C_n} \; q_{n,m}(\eta_{n,m}^{(k)} - 1)$
$\eta_{n,m'}^{(k+1)} := \eta_{n,m'}^{(k)} - 1$
end for

Algorithm 5.7.1 first solves the continuous problem (5.14) and then takes a finite number of steps (at most equal to the number of users in the cell) to find a discrete solution.

By setting $I_n(\boldsymbol{q}) = q_n^{\dagger}$, it can be shown that the function $\boldsymbol{I}(\boldsymbol{q}) = (I_n(\boldsymbol{q}))_{n \in \mathcal{N}}$ is standard [106] using exactly the same argument as used in Lemma 5.3.5. The standard power control algorithm can then be used to find the solution to (5.36).

5.8 Conclusions

Continuing our analysis of the flat fading resource allocation problem of Chapter 4, we investigated the impact of reducing the complexity of the original problem, by imposing additional constraints. We considered two such reduced complexity schemes. In both cases, the imposition of the additional constraints has the effect of almost halving the number of degrees of freedom available in the optimization problem.

The first reduced complexity scheme we considered was the "static bandwidth allocation" scheme (Scheme 2) which selects the bandwidth allocation to the users statically, and then adapts the transmit powers for the users in order to meet the normalized rate targets of the individual users. A distributed algorithm was developed to solve this problem.

The second reduced complexity scheme was the "uniform power spectral density (PSD) allocation" scheme (Scheme 3) which constrains each base station to use a uniform PSD across all users in its cell. Each base station can still choose an appropriate overall transmit power level and the bandwidth allocation to the users is dynamic. Two iterative algorithms were devised to solve this problem. The first of which was based on Yates framework [106], and it requires a nonlinear optimization to be solved using iterative methods within each step of the power control iteration. The second algorithm eliminates the need for the double iteration, by combining the steps of both iterations. As this algorithm does not fall under Yates framework, the convergence of this algorithm was proved by using the properties of the power vector sequence generated by the algorithm.

The performances of Scheme 1 (exact solution to the original problem), 2 and 3 were numerically evaluated via simulations. The numerical results showed that the loss in performance due to the loss in the number of degrees of freedom available to the optimization is small. Of the two reduced complexity schemes, Scheme 3 slightly outperforms Scheme 2.

We have also considered the discrete version of the uniform PSD allocation scheme and developed a distributed algorithm to solve it.

Chapter 6

Resource Allocation Under Frequency Selective Fading

6.1 Introduction

Chapters 4 and 5 consider the resource allocation problem for an OFDM cellular network in a static environment in which the channel gains do not change. In a multipath frequency selective fading environment, both signal and interference experience fluctuations, resulting in variations in the achieved normalized rate at the receiver. For the base station to have the perfect channel state information (CSI), the amount of feedback required and consequently the amount of signaling overhead incurred will be significant. Furthermore, if the channel changes quickly in time, the base station may not be able to maintain perfect CSI. Due to these reasons, the work in this chapter only assumes that the base stations have partial CSI, namely, the statistics of the fading.

In a fading environment in which the base stations do not have perfect CSI, a natural measure of performance is the outage probability. As before, we will assume that there is coding over the subchannels used by any particular user, and a user will be in outage if the total mutual information summed over the allocated subchannels falls short of the threshold needed to support the user's target data rate. The performance of any resource allocation algorithm can then be quantified by the probability of outage for each of the users.

Any one of the three deterministic schemes developed for the deterministic model in the preceding chapters could be used to obtain a resource allocation in the multipath frequency selective fading environment by simply replacing the random gains with their averages. However, it turns out that none of the three

schemes corresponding to the deterministic model has good performance with respect to the outage probability metric, at least when they are applied directly, replacing random gains by their averages.

Thus, we provide a layered approach to solve this outage probability based resource allocation problem. This two layer approach is presented in Section 6.2. The higher layer resource allocation problem deals with determining the transmit powers to be used by the base stations, and is solved with Scheme 3 (uniform transmit PSD allocation scheme) developed in Chapter 5 by considering the average gains of the links. The lower layer resource allocation problem is concerned with making the outage probabilities of the users "similar" by taking account of the statistical knowledge about the subchannels, and is formulated in Section 6.3. Section 6.4 solves the lower layer problem with the help of the central limit theorem. The performance of the two layer approach is numerically evaluated in Section 6.5 using Monte Carlo simulations. Section 6.6 investigates the effect of incorporating time diversity into the lower layer resource allocation problem.

6.2 Layered Approach

The approach of using any one of the deterministic schemes directly to obtain a resource allocation in a multipath frequency selective fading environment, by simply replacing the random gains by their averages will not lead to a good user performance with respect to the outage probability metric, as it has long been recognized that fade margins are required to account for statistical fluctuations in communication systems.

The fading margins can also be applied to the problem at hand. A fade margin typically refers to the allowance in transmit power to combat fading. Since the transmit powers are dynamically allocated in our context, we can apply a margin to the rate target of each user instead (which indirectly adds a margin to the transmit power), and thus it is referred to as a *rate margin*.

Applying any one of the three deterministic schemes by including a rate margin for each user to obtain a resource allocation will improve the outage perfor-

mance of the users. Ideally, appropriate rate margins should be applied to each user (the optimal margins will likely be different for different users) so that the users have similar outage probabilities (i.e. the measure of variance across the user outage probabilities is minimal). Unfortunately, it does not seem possible in our case to predetermine appropriate rate margin for each user individually; at least, it is beyond the scope of this thesis.

Therefore, we propose a two layer approach to solve the outage based resource allocation problem. The first (higher) layer deals with the allocation of the transmit powers to users by considering the average gains of the links. A fixed rate margin is added to the users' normalized rate requirements during the running of the deterministic algorithm. This in effect corresponds to targeting higher normalized rates for the users. However, as noted above, employing a rate margin during the higher layer allocation alone does not lead to a satisfactory user performance, if the measure of variance across the outage probability values of the users is significant. Some form of resource reallocation is required to reduce this variance.

The lower layer problem exactly addresses this issue. The specific problem associated with the lower layer resource allocation is the problem of "balancing" the outage probabilities (i.e., minimizing the maximum outage probability) among the users by reallocating the subchannels by taking account of the statistical knowledge about the subchannels.

Of the three deterministic schemes that we have studied, Scheme 3 provides the appropriate structure for the development of the lower layer resource allocation. By using Scheme 3 at the higher layer and fixing the transmit powers to the values obtained by the higher layer, the lower layer resource allocation problem can be solved independently in each cell, without knock-on effects between cells, and it avoids the need for iterative algorithms which may take too long to converge. Refer to Section 6.4.1 for a detailed discussion.

Recall that when Scheme 3 was solved in a flat fading environment in Chapter 5, explicit interference averaging (i.e., Latin square based fast frequency hopping) was not necessary since using the uniform PSD across all subcarriers was

sufficient to achieve the required interference averaging. However, in a frequency selective fading environment, the statistics of the subcarriers will be different if fast frequency hopping is not used together with Scheme 3. We will then be required to consider allocation of individual subcarriers to users. On the other hand, by employing fast frequency hopping with Scheme 3, we only need to model the number of subchannels for each user, as in the case under flat fading.

Scheme 3 is used with a rate margin to first allocate transmit powers to the cell sites by considering the average gains of the links across all subcarriers. The subcarrier allocations obtained in that scheme are then purely virtual, although they can be used as the input to the proposed lower layer subchannel allocation (LLSA) algorithm. Scheme 3 is implemented on a slow timescale to track large-scale ("macroscopic") changes in the network, such as changes in cell loadings (both from call arrivals and departures, and handover between cells), and large-scale mobility of users within a cell. On a faster timescale, subchannels are allocated between users within the cell using the LLSA algorithm which attempts to balance the outage probability using the knowledge about the statistics of the subchannels. The specific statistics we use in the optimization are the mean and the variance of the achieved normalized rate of a subchannel. Each receiver will measure these statistics over an integer number of hopping cycles and feed them back to its base station.

The proposed LLSA algorithm aims to obtain a subchannel allocation that minimizes the maximum outage probability among the users, given the total transmit power in each cell. Note here that since the number of subchannels is a discrete quantity it may not be possible to obtain an allocation that exactly equalizes the outage probabilities among all users, but if there are a large number of subcarriers, it will come fairly close.

The layered approach is summarized here. The system implements a Latin square design based fast frequency hopping to achieve interference averaging.

1. ***Higher Layer - Power Allocation:*** The transmit powers for the users are chosen with Scheme 3 (uniform transmit PSD allocation scheme). A rate margin is included during the running of the algorithm to account for the

statistical fluctuations. The subchannel allocation obtained during this step is virtual.

2. ***Lower Layer - Subchannel Reallocation:*** Fixing the transmit powers to the values obtained during the first step, a subchannel allocation for the users in each cell is determined using the Lower Layer Subchannel Allocation (LLSA) algorithm (to be developed in this chapter) to minimize the maximum outage probability of the users.

6.3 Lower Layer Resource Allocation

The lower layer resource allocation problem deals with "balancing" the outage probabilities of the users, i.e., minimizing the maximum outage probability of the users. The approach proposed in Section 6.2 enables us to solve the lower layer resource allocation problem locally in each cell independent of other cells.

Since each subchannel hops over every subcarrier during a hopping cycle, the subchannels on a particular link can be considered identically distributed. The number of symbol periods during which any two subchannels on a link are highly correlated will be small during a hopping cycle. Thus, if no user gets allocated a large proportion of the subchannels, the average correlation between subchannels on a given link will be very low, and can be neglected. This allows us to model a simplified version of the problem where the fading on the subchannels within a link are independent, and identically distributed (i.i.d.).

Let N_c be the number of subcarriers in the system, and consider a cell n. Let η_m be the number of subchannels allocated to user $m \in C_n$; since the problem of subchannel allocation in cell n does not depend on the subchannel allocations in any other cell (once the powers are fixed), we drop the label n from the notation at this point. The corresponding weight is $w_m = \eta_m / N_c$. Let $\boldsymbol{\eta} = (\eta_m)_{m \in C_n}$. The instantaneous gain of a subchannel i consists of two components: $\Gamma_{n,m}$ which is the average gain and is the same for all subchannels on that link, and $F_{n,m}^{(i)}$ which models fading on subchannel i. The instantaneous signal to interference

and noise ratio (SIR) on subchannel i is given by:

$$\gamma_{n,m}^{(i)} = \frac{\Gamma_{n,m} q_n F_{n,m}^{(i)}}{\sigma_m^2 + \sum_{k \neq n} \Gamma_{k,m} q_k F_{k,m}^{(i)}}. \tag{6.1}$$

The corresponding instantaneous normalized rate of subcarrier i is

$$c_m^{(i)} = \frac{1}{N_c} f(\gamma_{n,m}^{(i)}) \tag{6.2}$$

where $f(\gamma)$ is the spectral efficiency of the link (n, m) when the SIR is γ, and is characterized in Section 4.2. The achieved normalized rate of user m is the sum of the achieved normalized rates of individual subchannels allocated to that user: $\sum_{i=1}^{\eta_m} c_m^{(i)}$. The outage probability of user m with η_m subchannels is given by

$$P_m^{\text{out}}(\eta_m) = \mathbb{P}\left(\sum_{i=1}^{\eta_m} c_m^{(i)} < c_m^{tar} \right). \tag{6.3}$$

Note that $P_m^{out}(\eta_m)$ is a monotonically non-increasing function of η_m since allocating one more subchannel to a user will not increase that user's outage probability.

The optimization problem corresponding to minimizing the maximum outage probability of the users is:

$$\min_{\boldsymbol{\eta}} \max_{m \in C_n} P_m^{\text{out}}(\eta_m) \tag{6.4}$$

such that

$$\begin{aligned} \sum_{m \in C_n} \eta_m &= N_c, \\ \eta_m &\in \mathbb{Z}_+, \quad \forall m \in C_n. \end{aligned}$$

Solving (6.4) exactly will require the knowledge of the distribution of $c_m^{(i)}$'s. However, a good approximation can be made using only the mean and variance of the sum of random variables.

Let $Z_m = \sum_{i=1}^{\eta_m} c_m^{(i)}$, and mean and variance of $c_m^{(i)}$ be μ_m and β_m^2 respectively.

Define,

$$\hat{Z}_m(\eta_m) = \frac{Z_m - \eta_m \mu_m}{\sqrt{\eta_m}\,\beta_m}. \tag{6.5}$$

Then the outage probability of user m is given by

$$\hat{P}_m^{\text{out}}(\eta_m) \;=\; \mathbb{P}\left(\hat{Z}_m(\eta_m) < B_m(\eta_m)\right) \tag{6.6}$$

where

$$B_m(\eta_m) = \frac{c_m^{tar} - \eta_m \mu_m}{\sqrt{\eta_m}\,\beta_m}. \tag{6.7}$$

Since $B_m(x)$ is continuous and $dB_m/dx < 0$ for $x \in \mathbb{R}_+$, $B_m(\cdot)$ is monotonically decreasing in $\mathbb{R}_+$.

By the central limit theorem, the $\hat{Z}_m(\eta_m)$ are roughly identically distributed provided all users are allocated a sufficient number of subchannels (the η_m are large enough). Note that this approximation concerns the similarity in the distribution of the $\hat{Z}_m(\eta_m)$ to each other, and not their similarity to the Gaussian.

From the right hand side of (6.6), it is clear that minimizing the maximum of $\hat{P}_m^{\text{out}}(\eta_m)$, under the central limit theorem approximation, is roughly equivalent to minimizing the maximum of $B_m(\eta_m)$ among the users. Therefore, the optimization we need to solve is:

$$\min_{\boldsymbol{\eta}} \; \max_{m \in C_n} \; B_m(\eta_m) \tag{6.8}$$

such that

$$\begin{aligned} \sum_{m \in C_n} \eta_m \;&=\; N_c, \\ \eta_m \;&\in\; \mathbb{Z}_+, \qquad \forall m \in C_n. \end{aligned}$$

We solve Problem (6.8) exactly in two steps in the next section.

6.4 Lower Layer Subchannel Allocation (LLSA) Algorithm

Problem (6.8) is a nonlinear integer programming problem, which is combinatorial in nature. There may be multiple allocations solving it (existence of a solution is guaranteed as long as $N_c \geq L_n$, with L_n being the number of users in cell n). We now derive an $O(L_n)$ algorithm that solves (6.8). We will refer to the following development as Algorithm (LLSA).

It is easier to first solve the continuous relaxation of (6.8), by dropping the integrality constraints on the number of subchannels allocated to each user. We will then use the solution of this relaxed problem to obtain a solution to Problem (6.8) in a finite number of steps.

The continuous problem associated with (6.8) is

$$\min_{\boldsymbol{x}} \max_{m \in C_n} B_m(x_m) \tag{6.9}$$

such that

$$\begin{aligned} \sum_{m \in C_n} x_m &= N_c, \\ x_m &> 0, \qquad \forall m \in C_n, \end{aligned}$$

where $\boldsymbol{x} = (x_m)_{m \in C_n}$ and x_m are not required to be integers. The following lemma characterizes the properties of the solution of the relaxed problem (6.9).

Lemma 6.4.1. *The relaxed problem* (6.9) *has a unique solution* $\mathbf{x}^*$*, and there exists a* $\tilde{B}^*$ *such that,* $B_m(x_m^*) = \tilde{B}^*$ *for all* $m \in C_n$*.*

Proof. The inverse $x_m(\cdot)$ of $B_m(\cdot)$ satisfies $B_m(x_m(B)) = B$ for any $B \in \mathbb{R}$. The expression for $x_m(\cdot)$ can be obtained by solving the quadratic implied by (6.7), i.e., $\mu_m \, x_m(B) + \beta_m \, B \, \sqrt{x_m(B)} - c_m^{tar} = 0$:

$$x_m(B) = \left(\frac{\sqrt{(B\beta_m)^2 + 4\mu_m c_m^{tar}} - B\beta_m}{2\mu_m} \right)^2 \tag{6.10}$$

where the "+" has been taken since $\sqrt{x_m(B)} > 0$. Since $x_m = B_m^{-1}$, $x_m(\cdot)$ is monotonically decreasing. Since $x_m(\cdot)$ is monotonically decreasing and continuous, it follows that $\sum_{m\in C_n} x_m(B)$ is also. Therefore, $\sum_{m\in C_n} x_m(B) = N_c$ has a unique solution $B = \tilde{B}^*$ since $x_m(-\infty) = \infty$ and $x_m(\infty) = 0$. Thus the unique solution to (6.9) is $x_m^* = x_m(\tilde{B}^*)$ for all $m \in C_n$. Consequently, $B_m(x_m^*) = \tilde{B}^*$, for all $m \in C_n$. □

We note that $\tilde{B}^*$ is the minimum of the relaxed problem (6.9). Problem (6.9) can be solved by first determining the value of $\tilde{B}^*$ and then using it to determine $\boldsymbol{x}^*$. The value of $\tilde{B}^*$ can be determined by a bisection search [10]. This involves constructing bounds for $\tilde{B}^*$ first: an upper bound B^u and a lower bound B^l satisfying $\sum_{m\in C_n} x_m(B^l) \geq N_c$ and $\sum_{m\in C_n} x_m(B^u) \leq N_c$. The following lemma provides a simple way of choosing these bounds.

Lemma 6.4.2. *For an* $\mathbf{x} > 0$ *satisfying* $\sum_{m\in C_n} x_m = N_c$*, the following holds:*

$$\min_{m\in C_n} B_m(x_m) \;\leq\; \tilde{B}^* \;\leq\; \max_{m\in C_n} B_m(x_m). \tag{6.11}$$

Proof. Since $\sum_{m\in C_n} x_m = \sum_{m\in C_n} x_m^* = N_c$, there exists an m' such that $x_{m'} \leq x_{m'}^*$ and another m'' such that $x_{m''} \geq x_{m''}^*$. The required result follows from the fact that $B_m(\cdot)$ is monotonically decreasing in $\mathbb{R}_+$ and $B_m(x_m^*) = \tilde{B}^*$ for all $m \in C_n$. □

Let B^* be the minimum of the original problem (6.8). Then, $\tilde{B}^* \leq B^*$, i.e., $\tilde{B}^*$ is a lower bound on the minimum of Problem (6.8).

Now define a vector $\bar{\boldsymbol{\eta}} = (\bar{\eta}_m)_{m\in C_n}$ with:

$$\bar{\eta}_m = \lceil x_m(\tilde{B}^*) \rceil, \qquad \forall m \in C_n. \tag{6.12}$$

Since $\bar{\eta}_m \geq x_m(\tilde{B}^*)$ and $B_m(\cdot)$ is monotonically decreasing, $B_m(\bar{\eta}_m) \leq \tilde{B}^* \leq B^*$, $\forall m \in C_n$. Furthermore, $\sum_{m\in C_n} \bar{\eta}_m \geq N_c$.

If $\sum_{m\in C_n} \bar{\eta}_m = N_c$ then Problem (6.9) has an integral solution, which will also solve (6.8).

Turning our attention to the typical case that $\sum_{m\in C_n} \bar{\eta}_m \geq N_c + 1$, we note that $B_m(\bar{\eta}_m) \leq B^*$, $\forall m \in C_n$. Thus, we can apply the following lemma and reduce the

excess in the number of subchannels allocated.

Lemma 6.4.3. *For any* $\boldsymbol{\eta} = (\eta_m)_{m \in C_n}$, *if*

- $B_m(\eta_m) \leq B^*$, $\forall m \in C_n$, *and*
- $\sum_{m \in C_n} \eta_m \geq N_c + 1$,

then,

$$\min_{m \in C_n} B_m(\eta_m - 1) \quad \leq \quad B^*. \tag{6.13}$$

Proof. Let $\boldsymbol{\eta}^* = (\eta_m^*)_{m \in C_n}$ be a subchannel allocation (not necessarily unique) that solves Problem (6.8). Since $\sum_{m \in C_n} \eta_m \geq N_c + 1$ and $\sum_{m \in C_n} \eta_m^* = N_c$, there exists a user $m' \in C_n$ such that $\eta_{m'} > \eta_{m'}^*$ or equivalently $(\eta_{m'} - 1) \geq \eta_{m'}^*$. Since $B_m(\cdot)$ is monotonically decreasing, $B_{m'}(\eta_{m'} - 1) \leq B_{m'}(\eta_{m'}^*) \leq B^*$ where the second inequality follows from the fact that $B^* = \max_{m \in C_n} B_m(\eta_m^*)$. The required result follows. □

We now show how to construct an integral solution of Problem (6.8) by successively applying Lemma 6.4.3, starting with the allocations in $\bar{\boldsymbol{\eta}}$, and terminating after a finite number of steps in such a solution.

We start with $\boldsymbol{\eta}^{(0)} = \bar{\boldsymbol{\eta}}$ and note that this allocation involves $r = \sum_{m \in C_n} \bar{\eta}_m - N_c$ excess subchannels, which need to be removed. We do this iteratively.

Suppose $\boldsymbol{\eta}^{(k)}$ has $r - k > 0$ excess subchannels, and satisfies $B_m(\eta_m^{(k)}) \leq B^*$, for all $m \in C_n$. Then Lemma 6.4.3 applies. In particular, removing a subchannel from user m' that satisfies

$$m' = \arg \min_{m \in C_n} B_m(\eta_m^{(k)} - 1)$$

results in a new subchannel allocation $\boldsymbol{\eta}^{(k+1)}$ that has $r - (k+1)$ excess subchannels, and satisfies $B_m(\eta_m^{(k+1)}) \leq B^*$ for all $m \in C_n$. By induction, we obtain after r steps that there are no excess subchannels, so we must be at a solution to Problem (6.8).

	η_m						
m	1	2	3	4	5	6	7
1	3.5	2.6	**1.8**	1.1	**0.5**	-0.2	-0.8
2	3.1	**2.2**	1.3	0.6	-0.4	-1.2	-1.9
3	2.8	**1.9**	1.0	0.2	-0.3	-0.8	-1.1

Table 6.1: Values of $B_m(\eta_m)$ for a cell with 3 users and 7 subchannels

Note that the algorithm is not equivalent to substracting one subchannel each from the r users with the smallest $B_m(\bar{\eta}_m - 1)$, as there are cases where more than one subchannel must be removed from the same user. We demonstrate this with a simple numerical example. Consider a cell with $L_n = 3$ users and $N_c = 7$ subchannels. The $B_m(\cdot)$ values for the cell are given in Table 6.1. Note that the entry in the i^{th} column of the m^{th} row is the value of $B_m(i)$, the value of B_m when user m is allocated i subchannels. Suppose that the solution to the relaxed problem is $\boldsymbol{x} = [4.2, 1.6, 1.2]$, giving an initial allocation of $\bar{\boldsymbol{\eta}} = [5, 2, 2]$. The corresponding values of B_m are $[0.5, 2.2, 1.9]$. There are $r = 2$ excess subchannels to be removed. Applying the LLSA algorithm will result in 2 subchannels to be removed from user 1 resulting in an allocation $\boldsymbol{\eta}^* = [3, 2, 2]$. The corresponding values of B_m are $[1.8, 2.2, 1.9]$, with a maximum value of 2.2.

We have therefore provided an algorithm, namely the Lower Layer Subchannel Allocation (LLSA) algorithm, that finds an integral solution to the subchannel allocation problem. We first solve the relaxed problem, using the bisection method. We then take a finite number of steps (at most equal to the number of users in the cell) to find an integral solution.

6.4.1 Applicability of the Approach

Of our three deterministic schemes, it is only practical to apply the lower layer subchannel allocation algorithm to Scheme 3, because the central limit theorem approximation assumes that all subchannels have the same mean and variance, independent of the subchannel allocation, and this mean and variance needs to be measured. With Scheme 1 or Scheme 2, one can average the mean and vari-

ance across subchannels, but the values obtained are a function of the subchannel allocation itself. Furthermore, in Scheme 3, the total power allocated to the base station does not change as a function of the subchannel allocation, whereas it does change for Schemes 1 and 2. The theory applies to Scheme 3 precisely because the power spectral density is constant for that scheme.

6.5 Numerical Results

In this section, we numerically evaluate the performance of the Lower Layer Subchannel Allocation (LLSA) algorithm under a multipath frequency selective fading environment.

We use a simulation setup similar to that of Section 5.5 with a cellular network having 9 equal-sized square cells in an area of 3 km $\times$ 3 km. There are 90 users uniformly randomly located in the network. Each user selects its normalized rate target from a set of four available values. The path loss exponent is 3 (with the reference distance set to 500 m). The log-normal shadowing has a mean of 0 dB and a standard deviation of 8 dB for all distances. The noise spectral density is 10^{-19} W/Hz at each receiver. A network realization in this context refers to a random realization of user locations and shadowing.

There are 113 subcarriers in the system. As the FFT implementation requires the number of subcarriers to be a power of 2, this can correspond to a system with 128 subcarriers, with the remainder to be used as pilot and null subcarriers. Within each link, the channel gains of the subcarriers are independent and Rayleigh distributed about the flat fading gain (path loss plus log-normal shadowing).

The system uses explicit interference averaging. Fast frequency hopping patterns based on Latin square design were implemented for this purpose, using the simple construction (given N_c is prime) described in Section 3.7. The subchannel allocation to each user is discrete.

We refer to the scheme corresponding to our two layer approach as Scheme TL (two layer). We compare the outage performance of Scheme TL with that of the

following schemes: Schemes 1, 2 and 3. Schemes 1, 2 and 3 use the resource allocations determined by the deterministic algorithms. For Scheme TL, the cell level transmit powers are determined using Scheme 3. The LLSA algorithm is then applied to find the subchannel allocation that achieves "outage balancing", for the given cell transmit powers.

For Schemes 1, 2 and 3, discrete subchannel allocations are obtained by using a rounding procedure (which is described in Section 4.6.1) in which the continuous allocations are rounded down only for those users who can afford it. With Scheme 2, the discretization can be applied before determining the powers with the deterministic algorithm, as it uses a static bandwidth allocation. However, with Scheme 1 and 3, the discretization are performed only after the deterministic algorithms have been run as both use a dynamic bandwidth allocation. For Scheme 1, the powers are updated to work with the discrete subchannel allocation (by running an algorithm similar to that of Scheme 2).

For Scheme TL, the use of a rounding procedure is not necessary since continuous subchannel allocations obtained by the higher layer are not used. The lower layer obtains discrete subchannel allocations for users by running the LLSA algorithm.

We study the performance of the schemes for two fading scenarios which differ from each other on how quickly the channel gains change. The Fast fading scenario represents the case in which the coherence time of the channel is of the same order as the symbol periods, and thus the channel gains change on a faster timescale than that of a hopping cycle. In this case, it is necessary to measure the statistics over several hopping cycles, to get the ensemble statistics. On the other hand, in a slower (but still moderately fast) fading scenario in which subcarrier gains remain unchanged over several hopping cycles, the statistics can be measured over a hopping cycle and the receivers will be able to track the changing statistics.

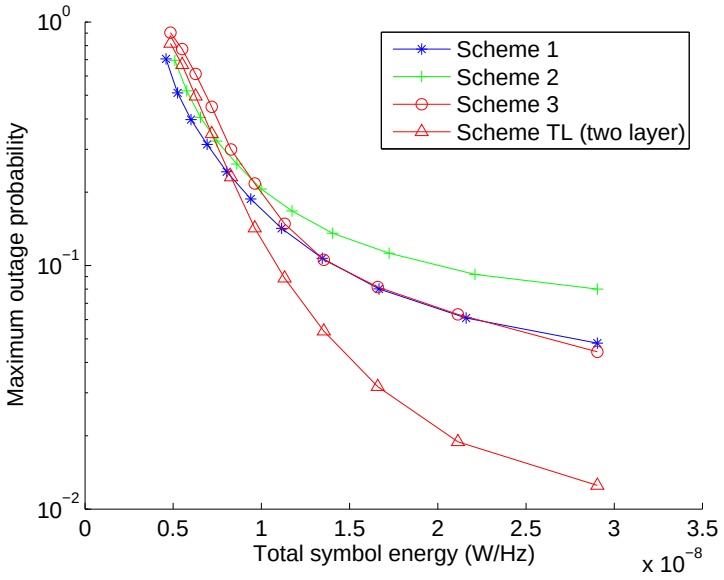


Figure 6.1: Fast fading: Mean over 50 network realizations of the maximum outage probability (over users). The total symbol energy increases with the rate margin applied.

6.5.1 Fast Fading

Fast fading scenario refers to the case in which the subcarrier gains change on the same timescale as that of the symbol periods. In this case, it is only possible to measure ensemble statistics. The statistics are computed by the receivers by averaging over the fast fading and the fast frequency hopping.

The outage performance of the schemes were computed for various rate margins in the range of 0% to 50% (a factor of 1 to 1.5). Monte Carlo simulations were used for calculating the achieved outage probabilities of the users. A given rate margin results in a different total symbol energy for each scheme. For each rate margin, 50 runs (i.e., computation of outages over 50 network realizations) were performed, and the mean value of the maximum user's outage is plotted against the mean total symbol energy in Figure 6.1. Firstly, the outage performances of all schemes are poor for low rate margins, and improve with the rate margin as expected. The performances of Scheme 1, 2 and 3 are similar. At low rate margins, the use of LLSA algorithm does not seem to improve the performance of Scheme TL. However, the performance of Scheme TL improves greatly as the rate margin is increased.

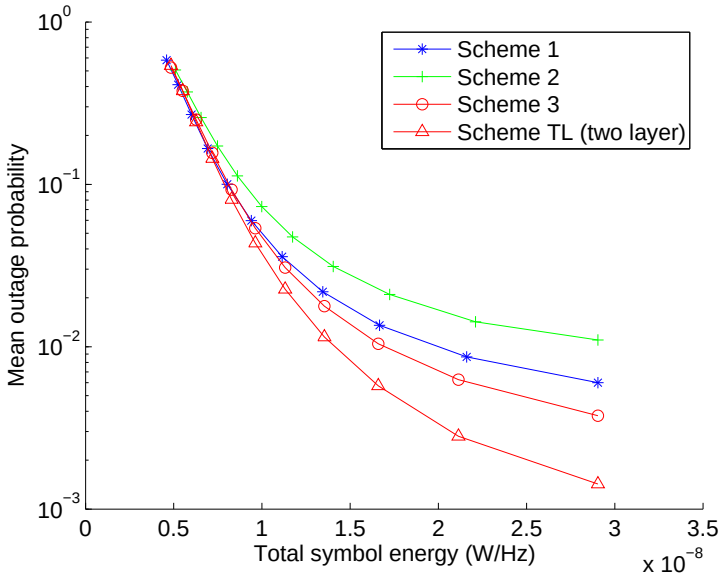


Figure 6.2: Fast fading: Mean over 50 network realizations of the mean outage probability (over users). The total symbol energy increases with the rate margin applied.

Similarly, Figure 6.2 plots the average of the mean achieved outage probability of the users as a function of mean total symbol energy. Scheme TL has better mean outage probabilities, even though we only targeted minimizing the maximum outage probability. This could be explained as follows. The subchannels are typically removed from users with large allocations ("donor" users) and given to users with small allocations ("recipient" users), since the latter have larger variance. The fractional change in the number of subchannels is larger for the "recipient" users, and so we expect their expected decrease in outage to be larger than the increase in outage of the "donor" users.

We compare the distribution of user outage probabilities for all schemes next. For a particular network realization, Figure 6.3 plots the outage probabilities of users grouped by cells, with the outage probabilities within each cell sorted in descending order. The rate margin for each scheme was chosen to make the total symbol energy for all schemes equal. From Figure 6.3, we can clearly observe that the LLSA algorithm indeed lessons the outage probability variations among the same cell users compared to the other three schemes. The corresponding maximum outage probabilities of the cells are plotted in Figure 6.4. From Figure 6.4,

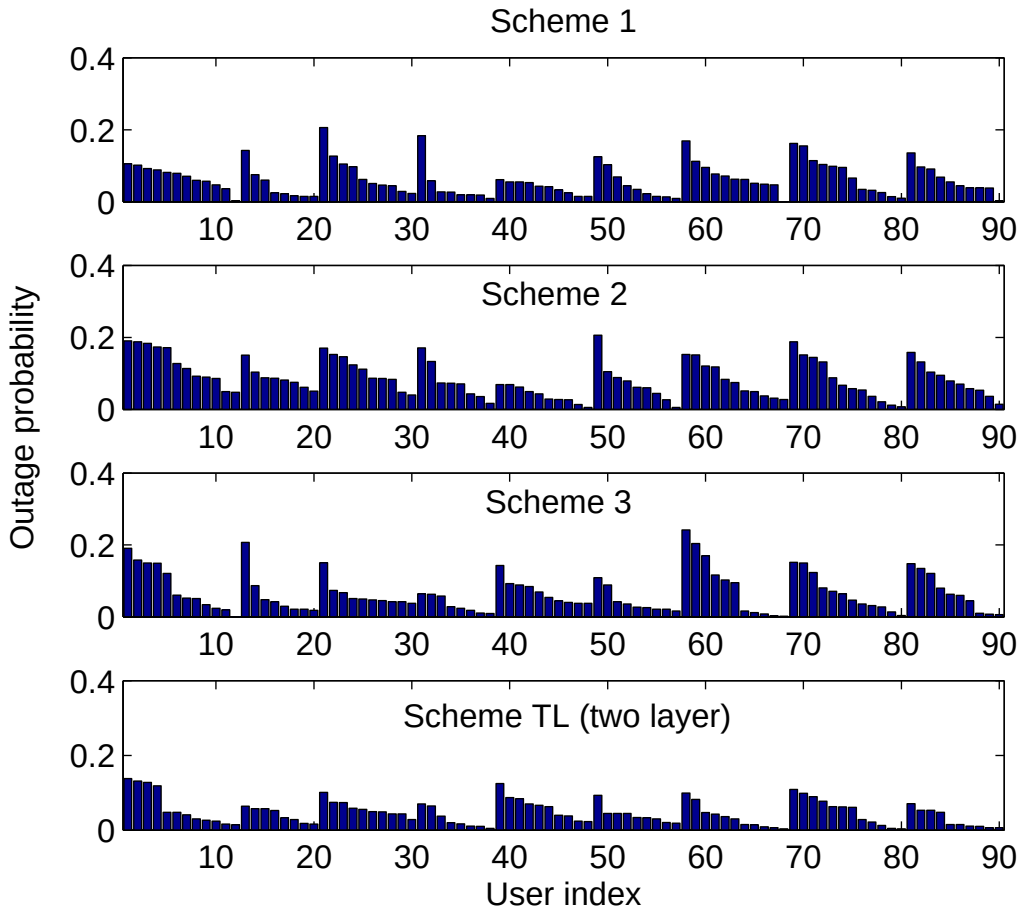


Figure 6.3: Fast fading: The outage probabilities of the users for a particular network realization. Rate margin for each scheme was selected to make all schemes use the same total symbol energy. The outage probabilities of same cell users are sorted in descending order.

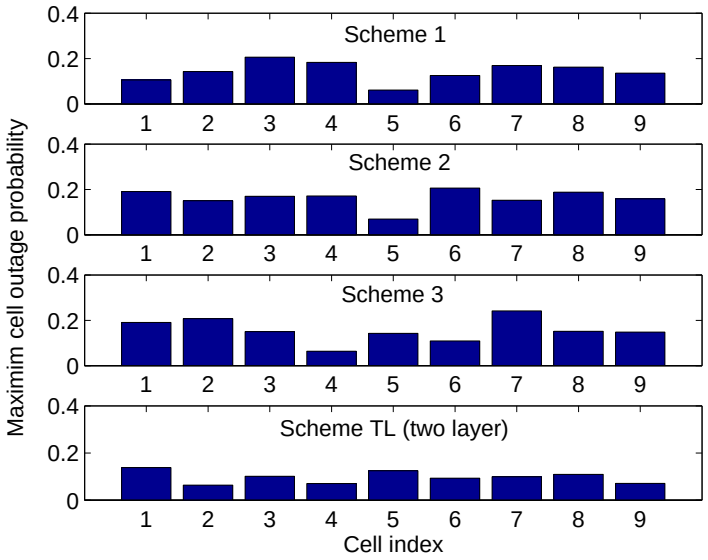


Figure 6.4: Fast fading: Maximum outage probability of the users in each cell (corresponding to Figure 6.3)

we can observe that for Scheme TL, the maximum outage probabilities are not dissimilar between cells, indicating that the lower layer approach of balancing the outages in each cell independently is a sensible one.

6.5.2 Slower Fading

The slower fading scenario refers to the case in which the subcarrier gains change on a slower timescale than that of a hopping cycle. The subcarrier gains remain unchanged over several hopping cycles, allowing the receivers to keep track of the gains of each channel realization. However, there may not be sufficient time for the base stations to acquire complete channel knowledge in order to perform a better optimization, and/or the amount of feedback information required for complete channel knowledge at the base stations is significant. For this reason, each receiver still feeds back only the statistics of the channel to the base station. The required statistics are computed by averaging over the fast frequency hopping for the current realization of the channel over a period of one hopping cycle.

The outage performance of the schemes under the influence of slower fading were computed for various rate margins in the range of 0% to 50%. For each rate

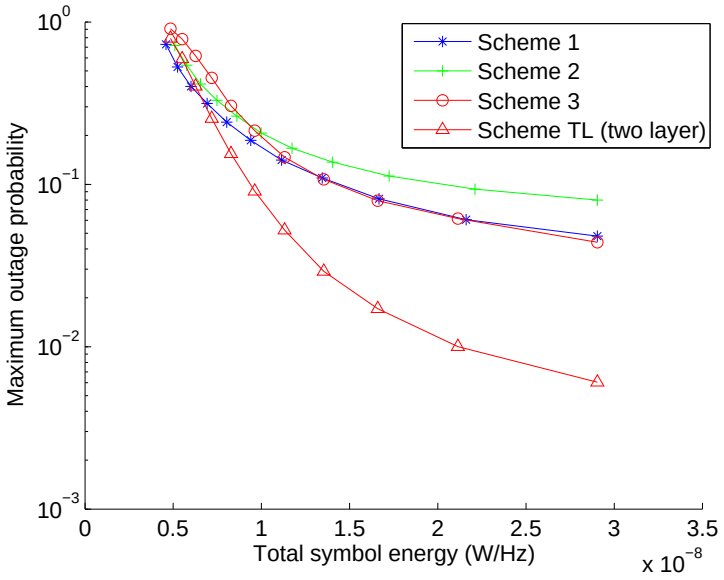


Figure 6.5: Slower fading: Mean over 50 network realizations of the maximum outage probability (over users). The total symbol energy increases with the rate margin applied.

margin, 50 runs (i.e., computation of outages over 50 network realizations) were performed. For each network realization, the outage probability values were averaged over 20 channel realizations. The resulting mean value of the maximum user's outage is plotted against the mean total symbol energy in Figure 6.5. The performance of the schemes are qualitatively same as that of the fast fading scenario. The major difference though is that Scheme TL has a better performance as compared to the fast fading scenario. This is to be expected as the lower layer subchannel allocation in the slower fading scenario is based on the statistics conditioned on the current channel realization.

Figure 6.6 plots the average of the mean achieved outage probability of the users as a function of mean total symbol energy. All schemes have the same qualitative behaviour as in the fast fading scenario. For a particular network realization, Figure 6.7 plots the outage probabilities of users grouped by cells, with the outage probabilities within each cell sorted in descending order while Figure 6.8 plots the corresponding maximum outage probability for each cell. The rate margin for each scheme was chosen to make the total symbol energy for all schemes equal. Figure 6.7 confirms that the LLSA algorithm achieves better

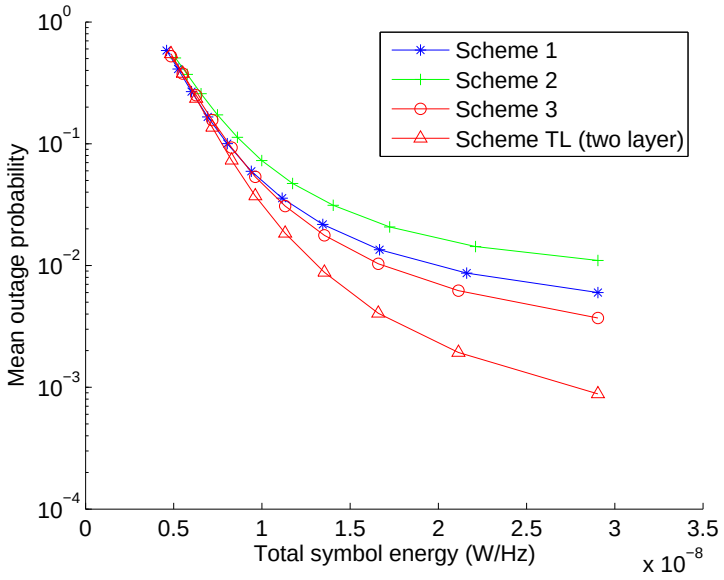


Figure 6.6: Slower fading: Mean over 50 network realizations of the mean outage probability (over users). The total symbol energy increases with the rate margin applied.

results with the slower fading scenario.

6.6 Incorporating Time Diversity

While studying the lower layer resource allocation problem in Section 6.3, we only considered the frequency diversity, i.e., there was coding over the subchannels, but not over the symbol periods. In this section, we investigate the case in which both time and frequency diversities are exploited. We note that this formulation is more applicable to a fast fading scenario where the subcarrier gains change on the same timescale as the symbol periods.

We formulate the lower layer resource allocation in this context with the same settings as that of Section 6.3, i.e., with the assumption that the average correlation between subchannels can be neglected, and consider the case that the fading across subchannels is independent, identically distributed (i.i.d.). Let T_D be the time diversity window (i.e., the number of symbol periods over which we code

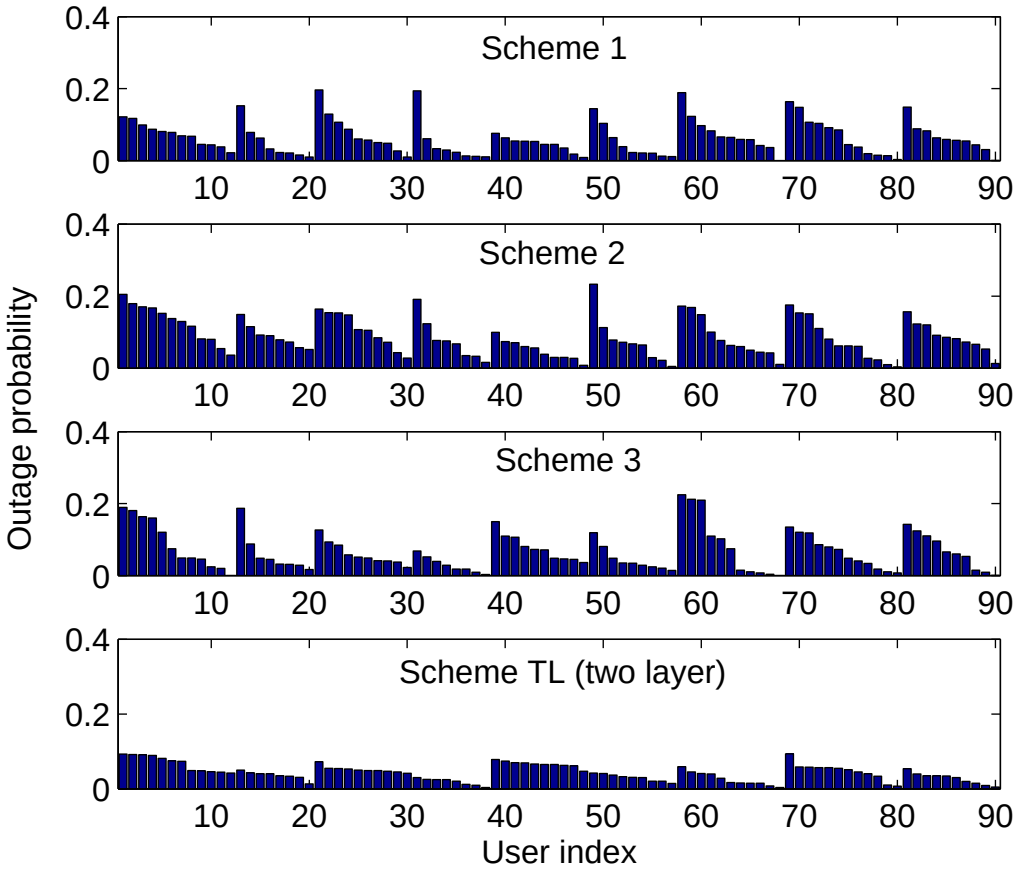


Figure 6.7: Slower fading: The outage probabilities of the users for a particular network realization. Rate margin for each scheme was selected to make all schemes use the same total symbol energy. The outage probabilities of same cell users are sorted in descending order.

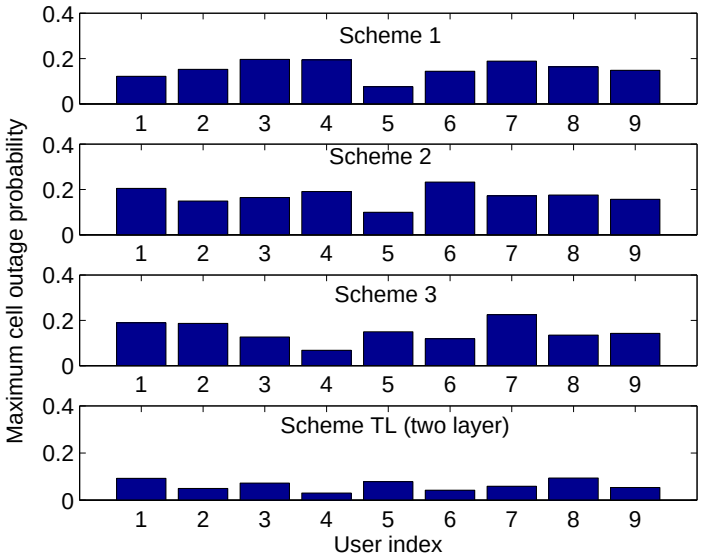


Figure 6.8: Slower fading: Maximum outage probability of the users in each cell (corresponding to Figure 6.7)

over). Then, we can write down the expression for the outage probability:

$$P_m^{\text{out,TD}}(\eta_m) \quad = \quad \mathbb{P}\left(\sum_{t=1}^{T_D}\sum_{i=1}^{\eta_m} c_m^{(i)} < T_D\, c_m^{tar}\right). \tag{6.14}$$

Note that an outage occurs when the achieved rate over a period of T_D symbol periods falls below $T_D\, c_m^{tar}$. Furthermore, within the time diversity window, the subchannels will occupy different sets of subcarriers at different symbol periods, as we consider a system that employs fast frequency hopping.

As the expression for the outage probability in (6.14) is numerically complex to evaluate, we make an approximation similar to that of Section 6.3. Let $\hat{Z}_m^{TD} = \sum_{t=1}^{T_D}\sum_{i=1}^{\eta_m} c_m^{(i)}$. Then, the outage probability can be approximated to:

$$\hat{P}_m^{\text{out,TD}}(\eta_m) \quad = \quad \mathbb{P}\left(\hat{Z}_m^{TD}(\eta_m) < B_m^{TD}(\eta_m)\right) \tag{6.15}$$

where

$$B_m^{TD}(\eta_m) \quad = \quad \frac{T_D\, c_m^{tar} - T_D\, \eta_m \mu_m}{\sqrt{T_D\, \eta_m}\, \beta_m}.$$

Thus, the optimization corresponding to outage balancing is:

$$\min_{\boldsymbol{\eta}} \max_{m \in C_n} B_m^{T_D}(\eta_m) \tag{6.16}$$

such that

$$\begin{aligned} \sum_{m \in C_n} \eta_m &= N_c, \\ \eta_m &\in \mathbb{Z}_+, \qquad \forall m \in C_n. \end{aligned}$$

We now show that the analysis of Section 6.4 can be directly applied to solve Problem (6.16). Rewriting $B_m^{TD}(\eta_m)$, we obtain,

$$B_m^{TD}(\eta_m) = \frac{c_m^{tar} - \eta_m \mu_m}{\sqrt{\eta_m}\,\frac{\beta_m}{\sqrt{T_D}}},$$

which is similar to $B_m(\eta_m)$ of (6.7), but with the variance being scaled down by a factor of T_D (or equivalently the standard deviation by a factor of $\sqrt{T_D}$). This scaling is due to the averaging effect of coding over a period of T_D symbol periods. Therefore, Problem (6.16) can be solved by using the LLSA algorithm just by replacing β_m with $\beta_m/\sqrt{T_D}$, and the analysis will remain the same.

Due to the time constraints, numerical studies have not been conducted to quantify the performance of the lower layer allocation technique considered in this section. We leave this as future work.

6.7 Conclusions

In this chapter, we considered the outage probability based resource allocation problem in a multiple user, multiple cell OFDM cellular network under multipath frequency selective fading. We emphasized the problems incurred with applying the deterministic algorithms developed for the flat fading environment directly and proposed a two layer approach as a solution.

The higher layer problem is used to obtain the transmit power allocation for

the users, and is solved by Scheme 3 using the average gains of the links and including a rate margin to account for the statistical fluctuations. The lower layer deals with the precise allocation of subchannels to users taking account of the statistical knowledge about the subchannels. The objective of the lower layer resource allocation is to "balance" the outages among the users, i.e., to minimize the maximum outage probability. Employing Scheme 3 at the higher layer and fixing the power levels to the values obtained during the higher layer allows the lower layer algorithm to be operated in each cell independently, without knock-on effects from one cell to another.

As the resuting lower layer problem is numerically complex, it was approximated using the central limit theorem. This approximate lower layer problem was solved exactly in two steps by the LLSA algorithm. The first is a continuous relaxation, and the second refines the relaxed solution to the true integer solution, in at most $L_n - 1$ steps, where L_n is the number of users in the cell.

The performance of the layered approach was numerically studied for two different fading scenarios: fast fading and slower fading. The simulation results show that the LLSA algorithm works very well in balancing the outage probabilities among the users in both scenarios. Furthermore, it was observed that the maximum outage probabilities across the cells are similar, indicating that most of the gains from the lower layer subchannel allocation come from a sensible allocation of subchannels within each cell. The LLSA algorithm exhibits a better performance for the slower fading scenario as the subchannel allocation obtained with the LLSA algorithm in this case is based on statistics conditioned on the current channel states.

We also showed that when time diversity is incorporated into the lower layer resource allocation problem, this problem can be solved by using the analysis corresponding to the lower layer resource allocation that does not use time diversity.

Chapter 7

Conclusions and Future Directions

7.1 Summary of the Work

The primary goal of this thesis was to solve the downlink resource allocation problem for an OFDM cellular system in a multipath frequency selective environment. In particular, the focus of the problem was on the OFDM systems that implement *interference averaging*. A two layer approach was adopted to solve this problem. The higher layer problem is equivalent to a resource allocation problem in a flat fading environment, which was the topic of Chapters 4 and 5. Chapter 6 solves the frequency selective fading environment resource allocation problem with a two layer approach, by addressing the lower layer resource allocation problem of allocating subchannels to the users in each cell, using the power levels found by the higher layer solution.

The major contributions of each part of the thesis are described in the following sections.

7.1.1 Resource Allocation under Flat Fading

The objective of the higher layer resource allocation problem is to determine the optimal resource allocation for the users that minimizes the aggregate transmit power in the network while meeting the rate requirements of all users. The problem explicitly models the inter-cell interference. By allowing continuous allocation of subchannels to users, this problem was solved with an iterative algorithm which can be executed by each base station in a distributed fashion. The convergence of this algorithm is guaranteed as it falls under Yates framework [106]. A corresponding discrete version of this problem (i.e., the subchannel allocation is

modeled as discrete) was also solved using a similar approach.

We also considered two reduced complexity schemes for the higher layer resource allocation. The reduction in complexity here refers to the reduction in the number of degrees of freedom available for the optimization and is achieved by imposing additional constraints on the original optimization problem. The first of such schemes is the "static bandwidth allocation" scheme which takes the traditional approach of allocating the subchannels (bandwidth) statically to users and then adapting the transmit powers to meet user rate targets. This problem was solved by an iterative algorithm which is distributed and is provably convergent.

The second reduced complexity scheme is the "uniform transmit power spectral density (PSD) allocation" scheme which constrains each base station to use a uniform transmit power spectral density across all users in its cell. The subchannel allocation to users is dynamic and each base station can still select an appropriate cell level transmit power. The uniform transmit PSD allocation scheme provides the appropriate structure for the development of the two layer approach for the frequency selective fading environment, and was the major focus of Chapter 5.

The properties of optimal solution of the uniform transmit PSD allocation scheme were characterized and a distributed algorithm was devised using Yates framework [106]. As this iterative algorithm involves iterations within iterations, another distributed algorithm which only requires a single iteration was developed. However, this new algorithm does not fall under Yates framework. Therefore, the convergence of this algorithm was proved using a novel technique which makes use of the properties of the sequence of the power vectors generated by the algorithm.

Numerical studies were conducted to quantify the loss in performance of the reduced complexity schemes due to the reduction in the number of degrees of freedom in the optimization. The numerical results have shown that the loss in performance of the reduced complexity schemes is marginal even though the number degrees of freedom available has halved (for both schemes). Furthermore, of the two reduced complexity schemes, the uniform transmit PSD alloca-

tion scheme slightly outperforms the static bandwidth allocation scheme.

The discrete version of the uniform transmit PSD allocation scheme was also solved with a distributed algorithm.

7.1.2 Resource Allocation under Frequency Selective Fading

The resource allocation problem in a frequency selective fading environment was solved by adopting a two layer approach. The outage probability was used as the performance metric. The higher layer resource allocation was solved with uniform transmit PSD allocation scheme using average gains of the links and including a rate margin to account for the statistical fluctuations. The aim of the lower layer resource allocation is to "balance" the outages among the users (i.e., minimize the maximum outage probability). Fixing the transmit powers to the values obtained by the higher layer allocation enabled us to consider this outage balancing problem in each cell independently without the need to keep track of the inter-cell interference during the lower layer resource allocation.

As the resulting combinatorial problem is numerically complex to solve, the central limit theorem was used to approximate it by another, simpler combinatorial problem. This approximate problem was solved exactly in two steps: first step involves solving the continuous relaxation; then the relaxed solution is refined to obtain a solution to the approximate problem in linear time in the number of users in the cell.

The simulation results have shown that the two layer approach indeed achieves its objective of balancing the outages among the same cell users. Furthermore, the maximum outage probabilities of the cells were similar, indicating that the most of the gains of the lower layer allocation come from sensible allocation of subchannels within each cell.

A key benefit of using the layered approach of resource allocation is that it avoids a combinatorial explosion as the number of users and subcarriers grows large. In particular, the higher layer problem is not combinatorial (it is a continuous relaxation of a combinatorial problem) and it converges exponentially fast, as

is typical for these sorts of power control algorithms. The lower layer subchannel allocation problem is combinatorial, but it was solved in linear time using its inherent properties.

7.2 Future Directions

In Chapter 6, the lower layer resource allocation problem was formulated under the premise that when fast frequency hopping is employed, the average correlations between subchannels will be small and can be neglected. This allowed us to use the central limit theorem to obtain an approximate problem which is numerically simpler to solve.

When the correlations between subchannels are included into the lower layer problem, then the classical central limit theorem may no longer be used for the purpose of simplifying the problem. It may be possible to use a form of the central limit theorem that handles correlated variables (e.g., see [17]) to approximate the problem. However, this approach (of modeling the correlations of subchannels) will require the receivers to compute additional statistics and feed them back to the base station, which will add to the existing signaling overhead.

Bibliography

[1] IEEE Std. 802.11g, Wireless LAN medium access control (MAC) and physical layer (PHY) specifications, amendment 4: Further higher data rate extension in the 2.4 GHz band, Jul. 2003.

[2] OFDM for mobile data communications. White Paper, Flarion Technologies, Inc., Mar. 2003.

[3] ETSI EN 300 744, V1.5.1 (2004-06), Digital video broadcasting (DVB); framing structure, channel coding and modulation for digital terrestrial television (DVB-T), 2004.

[4] ETSI EN 300 401, V1.4.1 (2006-01), Radio broadcasting systems; digital audio broadcasting (DAB) to mobile, portable and fixed receivers, 2006.

[5] IEEE Std. 802.16e, Air interface for fixed and mobile broadband wireless access systems, Amendment 2: Physical and medium access control layers for combined fixed and mobile operation in licenced band and corrigendum 1, Feb. 2006.

[6] I. F. Akyildiz, S. Mohanty, and J. Xie. A ubiquitous mobile communication architecture for next-generation heterogeneous wireless systems. *IEEE Commun. Mag.*, 43(6):S29–S36, Jun. 2005.

[7] N. Andgart, P. Odling, A. Johansson, and P. O. Borjesson. Designing tone reservation PAR reduction. *EURASIP Journal on Applied Signal Processing*, 2006.

[8] A. G. Armada. A simple multiuser bit loading algorithm for multiuser WLAN. In *Proc. IEEE Int. Conf. Commun. (ICC)*, volume 4, pages 1168–1171, Jun. 2001.

[9] J. Armstrong. New OFDM peak-to-average power reduction scheme. In *Proc. IEEE Veh. Technol. Conf. (VTC)*, volume 1, pages 756–760, May 2001.

[10] K. Atkinson. *Elementary Numerical Analysis*. John Wiley & Sons, 1985.

[11] R. W. Bauml, R. F. H. Fischer, and J. B. Huber. Reducing the peak-to-average power ratio of multicarrier modulation by selected mapping. *IEEE Electron. Letters*, 32(22):2056–2057, Oct. 1996.

[12] R. F. Bellaver. Wireless: from Marconi to Macaw. In *Proc. IEEE Int. Symp. Technol. and Society*, pages 197–200, Sep. 2000.

[13] D. P. Bertsekas. *Nonlinear Programming*. Athena Scientific, second edition, 1999.

[14] E. Bidet, D. Castelain, C. Joanblanq, and P. Senn. A fast single-chip implementation of 8192 complex points FFT. *IEEE J. Solid-State Circuits*, 30(3):300–305, Mar. 1995.

[15] J. A. C. Bingham. Multicarrier modulation for data transmission: an idea whose time has come. *IEEE Commun. Mag.*, 28(5):5–14, May 1990.

[16] J. Blau. Telephone TV. *IEEE Spectrum*, 42(6):16–17, Jun. 2005.

[17] M. Blum. On the central limit theorem for correlated random variables. In *Proc. IEEE*, volume 52, pages 308–309, Mar. 1964.

[18] H. Bolcskei. Blind estimation of symbol timing and carrier frequency offset in wireless OFDM systems. *IEEE Trans. Commun.*, 49(6):988–999, Jun. 2001.

[19] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.

[20] M. Breiling, S. H. Muller-Weinfurtner, and J. B. Huber. SLM peak-power reduction without explicit side information. *IEEE Commun. Letters*, 5(6):239–241, Jun. 2001.

[21] R. A. Brualdi. *Introductory Combinatorics*. Pearson Prentice Hall, fourth edition, 2004.

[22] R. W. Chang. Synthesis of band limited orthogonal signals for multichannel data transmission. Bell Systems Technol. J., volume 45, Dec. 1966.

[23] R. W. Chang and R. A. Gibby. A theoretical study of performance of an orthogonal multiplexing data transmission scheme. *IEEE Trans. Commun.*, 16(4):529–540, Aug. 1968.

[24] M. Chiang. Balancing transport and physical layers in wireless multihop networks: jointly optimal congestion control and power control. *IEEE J. Select. Areas Commun.*, 23(1):104–116, Jan. 2005.

[25] S. Coleri, M. Ergen, A. Puri, and A. Bahai. Channel estimation techniques based on pilot arrangement in OFDM systems. *IEEE Trans. Broadcast.*, 48(3):223–229, Sep. 2002.

[26] A. J. Coulson. Maximum likelihood synchronization for OFDM using a pilot symbol: algorithms. *IEEE J. Selected Areas Commun.*, 19(12):2486–2494, Dec. 2001.

[27] T. M. Cover and J. A. Thomas. *Information Theory, Inference, and Learning Algorithms*. Cambridge University Press, 2003.

[28] N. Damji and T. Le-Ngoc. Adaptive BCPM downlink resource allocation strategies for multiuser OFDM in cellular systems. In *Proc. IEEE Int. Conf. on Commun. and Broadband Networking (ICBN)*, volume 1, pages 422–428, Oct. 2005.

[29] N. Damji and T. Le-Ngoc. Dynamic downlink OFDM resource allocation with interference mitigation and macro diversity for multimedia services in wireless cellular systems. *IEEE Trans. Veh. Technol.*, 55(5):1555–1564, Sep. 2006.

[30] J. A. Davis and J. Jedwab. Peak-to-mean power control in OFDM, Golay complementary sequences, and Reed-Muller codes. *IEEE Trans. Inform. Theory*, 45(7):2397–2417, Nov. 1999.

[31] O. Edfors, M. Sandell, J.-J. van de Beek, S. K. Wilson, and P. O. Borjesson. OFDM channel estimation by singular value decomposition. *IEEE Trans. Commun.*, 46(7):931–939, Jul. 1998.

[32] R. Fantacci, F. Chiti, D. Marabissi, G. Mennuti, S. Morosi, and D. Tarchi. Perspectives for present and future CDMA-based communication systems. *IEEE Commun. Mag.*, 43(12):134–142, Feb. 2005.

[33] B. L. Floch, M. Alard, and C. Berrou. Coded orthogonal frequency division multiplex. In *Proc. IEEE*, volume 83, pages 982–996, Jun. 1995.

[34] G. J. Foschini and Z. Miljanic. A simple distributed autonomous power control algorithm and its convergence. *IEEE Trans. Veh. Technol.*, 42(4), Nov. 1993.

[35] S. A. Grandhi, R. Vijayan, and D. J. Goodman. Distributed power control in cellular radio systems. *IEEE Trans. Commun.*, 42:226–228, Feb.-Apr. 1994.

[36] S. I. Grossman. *Calculus*. Saunders College Publishing, fifth edition, 1992.

[37] M. Haardt and W. Mohr. The complete solution for third-generation wireless communications: two modes on air, one winning strategy. *IEEE Personal Commun.*, 7(6):18–24, Dec. 2000.

[38] Z. Han, Z. Ji, and K. J. R. Liu. Power minimization for multi-cell OFDM networks using distributed non-coorperative game approach. In *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, volume 6, pages 3742–3747, Nov./Dec. 2004.

[39] L. Hanzo, M. Munster, B. J. Choi, and T. Keller. *OFDM and MC-CDMA for Broadband Multi-User Communications, WLANs and Broadcasting*. John Wiley & Sons, 2003.

[40] M. Hata. Empirical formula for propagation loss in land mobile radio services. *IEEE Trans. Veh. Technol.*, 29(3):317–325, Aug. 1980.

[41] R. W. Heath and G. Giannakis. Exploiting input cyclostationarity for blind channel identification in OFDM systems. *IEEE Trans. Signal Processing*, 47(3):848–856, Mar. 1999.

[42] J. Heo, I. Cha, and K. Chang. A novel transmit power allocation algorithm combined with dynamic channel allocation in reuse partitioning-based OFDMA/FDD system. In *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2006.

[43] J. Jang and K. B. Lee. Transmit power adaptation for multiuser OFDM systems. *IEEE J. Selected Areas Commun.*, 21(2):171–178, Feb. 2003.

[44] I. Kim, H. L. Lee, B. Kim, and Y. H. Lee. On the use of linear programming for dynamic subchannel bit allocation in multiuser OFDM. In *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, volume 6, pages 3648–3652, Nov. 2001.

[45] Y. Kim, B. J. Jeong, J. Chung, C. S. Hwang, J. S. Ryu, K. H. Kim, and Y. K. Kim. Beyond 3G: visions, requirements, and enabling technologies. *IEEE Commun. Mag.*, 41(3):120–124, Mar. 2003.

[46] D. Kivanc, G. Li, and H. Liu. Computationally efficient bandwidth allocation and power control for OFDMA. *IEEE Trans. Wireless Commun.*, 2(6):1150–1158, Nov. 2003.

[47] D. Kivanc and H. Liu. Subcarrier allocation and power control for OFDMA. In *Proc. Asilomar Conference on Signals, Systems and Computers*, volume 1, pages 147–151, Oct./Nov. 2000.

[48] B. S. Krongold and D. L. Jones. PAR reduction in OFDM via active constellation extension. *IEEE Trans. Broadcast.*, 49(3):258–268, Sep. 2003.

[49] B. S. Krongold and D. L. Jones. An active-set approach for OFDM PAR reduction via tone reservation. *IEEE Trans. Signal Processing*, 52(2):495–509, Feb. 2004.

[50] G. Kulkarni, S. Adlakha, and M. Srivastava. Subcarrier allocation and bit loading algorithms for OFDMA-based wireless networks. *IEEE Trans. Mobile. Comput.*, 4(6):652–662, Nov./Dec. 2005.

[51] H. Kwon and B. G. Lee. Distributed resource allocation through noncooperative game approach in multi-cell OFDMA systems. In *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2006.

[52] J. Lee, R. V. Sonalkar, and J. M. Cioffi. Multiuser bit loading for multicarrier systems. *IEEE Trans. Commun.*, 54(7):1170–1174, Jul. 2006.

[53] W. C. Y. Lee. *Mobile Cellular Telecommunications: Analog and Digital Systems*. MaGraw-Hill, second edition, 1995.

[54] J. Lei and T. S. Ng. A consistent OFDM carrier frequency offset estimator based on distinctively spaced pilot tones. *IEEE Trans. Wireless Commun.*, 3(2):588–599, Mar. 2004.

[55] K. K. Leung, C. W. Sung, W. S. Wong, and T. M. Lok. Convergence theorem for a general class of power-control algorithms. *IEEE Trans. Commun.*, 52(9):1566–1574, Sep. 2004.

[56] G. Li and H. Liu. Downlink dynamic resource allocation for multi-cell OFDMA system. In *Proc. IEEE Veh. Technol. Conf. (VTC)*, volume 3, pages 1698–1702, Oct. 2003.

[57] X. Li and L. J. Cimini Jr. Effects of clipping and filtering on the performance of OFDM. *IEEE Commun. Letters*, 2(5):131–133, May 1998.

[58] Z. Li, G. Zhu, W. Wang, and J. Song. Improved algorithm of multiuser dynamic subcarrier allocation in OFDM system. In *Proc. IEEE Int. Conf. Commun. Technol. (ICCT)*, volume 2, pages 1144–1147, Apr. 2003.

[59] D. J. Love and R. W. Heath. OFDM power loading using limited feedback. *IEEE Trans. Veh. Technol.*, 54(5):1773–1780, Sep. 2005.

[60] C. Mohanram and S. Bhashyam. A sub-optimal joint subcarrier and power allocation algorithm for multiuser OFDM. *IEEE Commun. Letters*, 9(8):685–687, Aug. 2005.

[61] B. Muquet, M. de Courville, and P. Duhamel. Subspace-based blind and semi-blind channel estimation for OFDM systems. *IEEE Trans. Signal Processing*, 50(7):1699–1712, Jul. 2002.

[62] R. Negi and J. M. Cioffi. Blind OFDM symbol synchronization in ISI channels. *IEEE Trans. Commun.*, 50(9):1525–1534, Sep. 2002.

[63] H. Nikookar and R. Prasad. On the sensitivity of multicarrier transmission over multipath channels to phase noise and frequency offset. In *Proc. IEEE Int. Symp. Pers., Indoor and Mobile Radio Commun. (PIMRC)*, volume 1, pages 68–72, Oct. 1996.

[64] H. Ochiai and H. Imai. Performance of the deliberate clipping with adaptive symbol selection for strictly band-limited OFDM systems. *IEEE J. Select. Areas Commun.*, 18(11):2270–2277, Nov. 2000.

[65] S. Ohmori, Y. Yamao, and N. Nakajima. The future generations of mobile communications based on broadband access technologies. *IEEE Commun. Mag.*, 38(12):134–142, Dec. 2000.

[66] T. Okumura, E. Ohmori, and K. Fukuda. Field strength and its variability in VHF and UHF land mobile service. Review Electrical Communication Laboratory, volume 16, number 9-10, Sep.-Oct. 1968.

[67] R. O'Neill and L. B. Lopes. Envelope variations and spectral splatter in clipped multicarrier signals. In *Proc. IEEE Int. Symp. Pers., Indoor and Mobile Radio Commun. (PIMRC)*, volume 1, pages 71–75, Sep. 1995.

[68] J. Papandriopoulos, S. Dey, and J. Evans. Distributed cross-layer design of MANETs in composite fading through global optimization. Submitted to IEEE J. Select. Areas Commun. URL: http://jpap.yi.org/research.html.

[69] J. Papandriopoulos, J. Evans, and S. Dey. Iterative power control and multi-user detection with outage probability constraints. In *Proc. IEEE Int. Conf. Commun.*, volume 4, pages 2509–2513, May 2003.

[70] K. Paterson. Generalized Reed-Muller codes and power control in OFDM modulation. *IEEE Trans. Inform. Theory*, 46(1):104–120, Jan. 2000.

[71] M. Pauli and H. P. Kuchenbecker. Minimization of the intermodulation distortion of a nonlinearly amplified OFDM signal. *Wireless Personal Communications*, 4(1):93–101, Jan. 1997.

[72] A. Peled and A. Ruiz. Frequency domain data transmission using reduced computional complexity algorithms. In *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Processing (ICASSP)*, volume 5, pages 964–967, Apr. 1980.

[73] A. Petropulu, R. Zhang, and R. Lin. Blind OFDM channel estimation through simple linear precoding. *IEEE Trans. Wireless Commun.*, 3(3):647–655, Mar. 2004.

[74] S. Pietrzyk and J. M. Jansson. Radio resource allocation for cellular networks based on OFDMA with QoS guarantees. In *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, volume 4, pages 2694–2699, Nov./Dec. 2004.

[75] T. Pollet and M. Moeneclaey. Synchronizability of OFDM signals. In *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, volume 3, pages 2054–2058, Nov. 1995.

[76] T. Pollet, M. van Bladel, and M. Moeneclaey. BER sensitivity of OFDM systems to carrier frequency offset and Weiner phase noise. *IEEE Trans. Commun.*, 43(2/3/4):191–193, Feb./Mar./Apr. 1995.

[77] G. J. Pottie and A. R. Calderbank. Channel coding strategies for cellular radio. *IEEE Trans. Veh. Technol.*, 44(4):763–770, Nov. 1995.

[78] J. G. Proakis. *Digital Communications*. McGraw-Hill, fourth edition, 2001.

[79] S. U. H. Qureshi. Adaptive equalization. In *Proc. IEEE*, volume 73, pages 1349–1387, Sep. 1985.

[80] T. S. Rappaport. *Wireless Communications*. Prentice Hall, second edition, 2002.

[81] W. Rhee and J. M. Cioffi. Increase in capacity of multiuser OFDM system using dynamic subchannel allocation. In *Proc. IEEE Veh. Technol. Conf. (VTC)*, volume 2, pages 1085–1089, May 2000.

[82] S. Roy and C. Li. A subspace blind channel estimation method for OFDM systems without cyclic prefix. *IEEE Trans. Wireless Commun.*, 1(4):572–579, Oct. 2002.

[83] M. Russell and G. L. Stuber. Interchannel interference analysis of OFDM in a mobile environment. In *Proc. IEEE Veh. Technol. Conf. (VTC)*, volume 2, pages 820–824, Jul. 1995.

[84] B. R. Saltzberg. Performance of an efficient parallel data transmission system. *IEEE Trans. Commun. Technol.*, 15:805–811, Dec. 1967.

[85] Z. Shen, J. G. Andrews, and B. L. Evans. Adaptive resource allocation in multiuser OFDM systems with proportional rate constraints. *IEEE Trans. Wireless Commun.*, 4(6):2726–2737, Nov. 2005.

[86] O. Simeone, Y. Bar-Ness, and U. Spagnolini. Pilot-based channel estimation for OFDM systems by tracking the delay-subspace. *IEEE Trans. Wireless Commun.*, 3(1):315–325, Jan. 2004.

[87] Z. Song, K. Zhang, and Y. L. Guan. Statistical adaptive modulation for QAM-OFDM systems. In *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, volume 1, pages 706–710, Nov. 2002.

[88] T. Starr, J. M. Cioffi, and P. Silverman. *Understanding Digital Subscriber Line Technology*. Prentice Hall PTR, 1999.

[89] C. Suh, Y. Cho, and S. Yoon. Dynamic subchannel and bit allocation in multiuser OFDM with a priority user. In *Proc. Int. Symp. Spread Spectrum Techniques and Applicat. (ISSSTP)*, pages 919–923, Aug./Sep. 2004.

[90] H. A. Taha. *Integer Programming - Theory, Applications, and Computations*. Academic Press, 1975.

[91] A. Tarighat and A. H. Sayed. An optimum OFDM receiver exploiting cyclic prefix for improved data estimation. In *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Processing (ICASSP)*, volume 4, pages 217–220, Apr. 2003.

[92] T. Thanabalasingham, S. V. Hanly, and L. L. H. Andrew. Dynamic allocation of subcarriers and transmit powers in a multiuser OFDM cellular network. *To be submitted to Trans. Wireless Commun.*

[93] T. Thanabalasingham, S. V. Hanly, and L. L. H. Andrew. Joint allocation of transmit power levels and degrees of freedom to links in a wireless network. In *Proc. Australian Communications Theory Workshop (AusCTW)*, pages 242–248, Feb. 2005. Available at IEEE Xplore.

[94] T. Thanabalasingham, S. V. Hanly, L. L. H. Andrew, and J. Papandriopoulos. Joint allocation of subcarriers and transmit powers in a multiuser OFDM cellular network. In *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2006.

[95] L. Tong and S. Perreau. Multichannel blind identification: From subspace to maximum likelihood methods. In *Proc. IEEE*, volume 86, pages 1951–1968, Oct. 1998.

[96] D. Tse and P. Viswanath. *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.

[97] J.-J. van de Beek, M. Sandell, and P. O. Borjesson. ML estimation of time and frequency offset in OFDM systems. *IEEE Trans. Signal Processing*, 45(7):1800–1805, Jul. 1997.

[98] R. van Nee and R. Prasad. *OFDM for Wireless Multimedia Communications*. Artech House, 2000.

[99] M. Wahlqvist, H. Olofsson, M. Ericson, C. Ostberg, and R. Larsson. Capacity comparison of an OFDM based multiple access system using different dynamic resource allocation. In *Proc. IEEE Veh. Technol. Conf. (VTC)*, volume 3, pages 1664–1668, May 1997.

[100] S. B. Weinstein and P. M. Ebert. Data transmission by frequency-division multiplexing using the discrete fourier transform. *IEEE Trans. Commun. Technol.*, 19(5):628–634, Oct. 1971.

[101] C. Y. Wong, R. S. Cheng, K. B. Lataief, and R. D. Murch. Multiuser OFDM with adaptive subcarrier, bit, and power allocation. *IEEE J. Select. Areas Commun.*, 17(10):1747–1758, Oct. 1999.

[102] C. Y. Wong, C. Y. Tsui, R. S. Cheng, and K. B. Letaief. A real-time sub-carrier allocation scheme for multiple access downlink OFDM transmission. In *Proc. IEEE Veh. Technol. Conf. (VTC)*, volume 2, pages 1124–1128, Sep. 1999.

[103] Y. Wu and W. Y. Zou. Orthogonal frequency division multiplexing: A multi-carrier modulation scheme. *IEEE Trans. Comsumer Electron.*, 41(3):392–399, Aug. 1995.

[104] A. D. Wyner. Multi-tone multiple access for cellular systems. AT&T Bell Labs. Technical Memorandum, BL011217-920812-12TM, 1992.

[105] Y. Yao and G. B. Giannakis. Rate-maximizing power allocation in OFDM based on partial channel knowledge. *IEEE Trans. Wireless Commun.*, 4(3):1073–1083, May 2005.

[106] R. D. Yates. A framework for uplink power control in cellular radio systems. *IEEE J. Select. Areas Commun.*, 13(7):1341–1347, Sep. 1995.

[107] W. Yu, W. Rhee, S. Boyd, and J. M. Cioffi. Iterative water-filling for Gaussian vector multiple-access channels. *IEEE Trans. Info. Theory*, 50(1):145–152, Jan. 2004.

[108] J. Zander. Distributed cochannel interference control in cellular radio systems. *IEEE Trans. Veh. Technol.*, 41(3):305–311, Aug. 1992.

[109] Y. J. Zhang and K. B. Letaief. Multiuser adaptive subcarrier-and-bit allocation with adaptive cell selection for OFDM systems. *IEEE Trans. Wireless Commun.*, 3(5):1566–1575, Sep. 2004.

[110] Z. Zhang, K. Long, M. Zhao, and Y. Liu. Joint frame synchronization and frequency offset estimation in OFDM systems. *IEEE Trans. Broadcast.*, 51(3):389–394, Sep. 2005.

[111] M. S. Zimmerman and A. L. Kirsch. The AN/GSC-10 (KATHRYN) variable rate data modem for HF radio. *IEEE Trans. Commun. Technol.*, 15(2):197–204, Apr. 1967.